Who is using UK Biobank and what are they doing?

Naomi Allen
Senior Epidemiologist, UK Biobank
Naomi.allen@ndph.ox.ac.uk
• Open access resource, available for bona fide researchers to conduct health-related research this is in the public interest

• Available for use by academia and commercial companies, both in the UK and abroad

• No preferential or exclusive access to the resource

• Researchers are obliged to return their results to UK Biobank so they can be shared with others
Research scope

• Research applications can:
  • be broad - do not need to focus on specific risk factors or health outcomes
  • involve a range of genotypes/phenotypes (i.e. GWAS/PheWAS)
  • be purely methodological

The main consideration is that the scope of the research project can be objectively defined, i.e. the intended research goals are clear
Access to UK Biobank so far

- Opened for access March 2012
- 7,500 approved registrations
- 1,400 applications submitted
  - 750 projects underway
- >400 publications
  - Projects cover a broad range of topics
  - See [www.ukbiobank.ac.uk/approved-research](http://www.ukbiobank.ac.uk/approved-research) for summaries
Number of registered researchers

7,500 registrations

UK | International
---|---
2012 | 15% | 15%
2013 | 15% | 15%
2014 | 26% | 44%
2015 | 44% | 54%
2016 | 54% | 68%
2017 | 68% |
Number of researchers worldwide

- United States of America: 2464
- Ireland: 3268
- United Kingdom (UK): 1015
- Canada: 439
- Brazil: 324
- Other countries with smaller numbers include:
  - 25: Brazil, Bolivia, Paraguay
475 original research papers (>3,700 citations)
<table>
<thead>
<tr>
<th>Author</th>
<th>Journal</th>
<th>Title</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global BMI Mortality Collaboration</td>
<td>Lancet (2016)</td>
<td>BMI and mortality (meta-analysis)</td>
<td>180</td>
</tr>
<tr>
<td>Okbay (SSGAC)</td>
<td>Nature (2016)</td>
<td>GWAS of education</td>
<td>159</td>
</tr>
<tr>
<td>Di Angelantonio (ERFC)</td>
<td>JAMA (2015)</td>
<td>Cardiometabolic multi-morbidity and mortality (meta-analysis)</td>
<td>105</td>
</tr>
<tr>
<td>Okbay (SSGAC)</td>
<td>Nat Genet (2016)</td>
<td>GWAS of subjective well-being and mental health</td>
<td>99</td>
</tr>
<tr>
<td>Miller (Oxford)</td>
<td>Nat Neurosci (2016)</td>
<td>Brain imaging</td>
<td>83</td>
</tr>
<tr>
<td>Ganna (Karolinska)</td>
<td>Lancet (2015)</td>
<td>Predictors of 5 year mortality</td>
<td>81</td>
</tr>
<tr>
<td>Wain (Leicester)</td>
<td>Lancet Resp Med (2015)</td>
<td>GWAS of smoking, lung function, and COPD (UK BiLEVE)</td>
<td>73</td>
</tr>
<tr>
<td>Hagenaars (Edinburgh)</td>
<td>Mol Psychiatry (2016)</td>
<td>Genetic factors in cognitive function and physical and mental health</td>
<td>62</td>
</tr>
</tbody>
</table>
750 institutions have published using UK Biobank data

<table>
<thead>
<tr>
<th>Name of Institution</th>
<th>Country</th>
<th>Number of publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Oxford</td>
<td>UK</td>
<td>104</td>
</tr>
<tr>
<td>University of Edinburgh</td>
<td>UK</td>
<td>78</td>
</tr>
<tr>
<td>University of Glasgow</td>
<td>UK</td>
<td>66</td>
</tr>
<tr>
<td>Broad Institute</td>
<td>USA</td>
<td>50</td>
</tr>
<tr>
<td>Imperial College London</td>
<td>UK</td>
<td>49</td>
</tr>
<tr>
<td>University of Cambridge</td>
<td>UK</td>
<td>48</td>
</tr>
<tr>
<td>Harvard University</td>
<td>USA</td>
<td>47</td>
</tr>
<tr>
<td>University of Bristol</td>
<td>UK</td>
<td>47</td>
</tr>
<tr>
<td>King's College London</td>
<td>UK</td>
<td>44</td>
</tr>
<tr>
<td>University College London</td>
<td>UK</td>
<td>43</td>
</tr>
</tbody>
</table>
Most productive non-UK institutions

625 (out of 750) institutions are non-UK

<table>
<thead>
<tr>
<th>Name</th>
<th>Country</th>
<th>Number of publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad Institute</td>
<td>USA</td>
<td>50</td>
</tr>
<tr>
<td>Harvard University</td>
<td>USA</td>
<td>47</td>
</tr>
<tr>
<td>University of Queensland</td>
<td>Australia</td>
<td>35</td>
</tr>
<tr>
<td>Massachusetts General Hospital</td>
<td>USA</td>
<td>27</td>
</tr>
<tr>
<td>Erasmus University Medical Center</td>
<td>Netherlands</td>
<td>26</td>
</tr>
<tr>
<td>Karolinska Institute</td>
<td>Sweden</td>
<td>25</td>
</tr>
<tr>
<td>University of Copenhagen</td>
<td>Denmark</td>
<td>23</td>
</tr>
<tr>
<td>University Medical Center Groningen</td>
<td>Netherlands</td>
<td>21</td>
</tr>
<tr>
<td>Uppsala University</td>
<td>Sweden</td>
<td>21</td>
</tr>
<tr>
<td>Stanford University</td>
<td>USA</td>
<td>21</td>
</tr>
</tbody>
</table>
Most common collaborations
Most popular areas of research

Number of publications

Public Health and Health Services
Genetics
Clinical Sciences
Psychology
Cardiovascular Medicine and Haematology
Ophthalmology and Optometry
Paediatrics and Reproductive Medicine
Artificial Intelligence and Image Processing
Neurosciences
Immunology
Statistics
Oncology and Carcinogenesis
Nutrition and Dietetics
Information Systems
Medical Microbiology
Other Studies in Human Society
Food Sciences
Computation Theory and Mathematics
Applied Economics
Medical Biochemistry and Metabolomics
Most popular health outcomes published

- Cardiovascular: 22%
- Metabolic and Endocrine: 14%
- Cancer: 14%
- Mental Health: 12%
- Stroke: 12%
- Generic Health Relevance: 10%
- Respiratory: 8%
- Oral and Gastrointestinal: 8%
- Eye: 5%
- Reproductive Health and Childbirth: 5%
- Musculoskeletal: 5%
- Inflammatory and Immune System: 5%
- Neurological: 4%
- Ear: 4%
- Infection: 2%
- Skin: 2%

Number of publications
What research is being done?

- Genetics
- Obesity
- Cancer
- Cardiovascular mortality
- Vision
- Physical activity
- Imaging
- Sleep
- Smoking
- Dementia
- Cognition
- Diet
- Biomarkers
- Hearing
- Inflammation
- Depression
- Osteoporosis
- COPD
- Multiple sclerosis
- Chronic's disease
- Hypertension
- Stroke
- Diabetes
- Ageing
- Reproduction
- Metabolic traits
- Haematology
- Social isolation
- Allergy
- Liver disease
- Oral health
- Pain
- Renal
- Reproduction
- Haematology
- Mendelian randomisation
- Built environment
- Primary care
- Socio-economic factors
- Multi-morbidity
- Gene-environment
- Air pollution
- Methodology
- Personalised medicine
- Brain
- Prediction