Parliamentary Office of Science and Technology

postnote

July 2002 Number 180

THE UK BIOBANK

In June 1999, the Wellcome Trust and Medical Research Council (MRC) announced a UK Population Biomedical Collection. This project – now called the UK Biobank – will establish a UK genetic databank by collecting blood samples from some 500,000 people. Researchers will apply for access to the Biobank data to study the factors behind common serious disorders such as heart disease, stroke, cancer and diabetes. This briefing describes the background to the UK Biobank, and examines issues raised such as consent, confidentiality, ownership and oversight.

Background

Many of the common 'killer diseases' result from a complex mixture of environmental and genetic factors:

- environmental – exposure to tobacco smoke, pollution, viruses, etc.; lifestyle factors such as diet and exercise; physiological risk factors such as blood pressure; demographic, medical and reproductive factors;
- genetic - the recent publication of a first draft of the human genome has seen a proliferation of hypotheses implicating genetic factors in a wide range of diseases.

In June 1999, the MRC and Wellcome Trust decided to set up a research resource to allow the individual and combined contribution of these factors to be assessed. The idea was to collect genetic and environmental (including lifestyle and medical) information from a large number of people and make this available to researchers studying the causes of common diseases of adult life. A working group published a draft protocol for the project in February 2002. Small-scale consultation exercises identified ethical issues likely to arise – these have been incorporated into the draft project design.

Proposed design

Size

Biobank will recruit at least 500,000 participants, and follow them up periodically over the next 10 years. Participants will be aged 45-69, since this will maximise the number of illnesses and deaths recorded. Around 1,000 deaths or illnesses for each disease over a 10 year follow up period should allow studies to tease out the effects of genetic and environmental factors. Over 10 years the cohort chosen should deliver enough cases to allow the study of a wide range of cancers (including stomach, ovarian, bladder, non-Hodgkins lymphoma, lung, prostate, colorectal and breast), heart attacks, heart disease, stroke, diabetes, hip fracture, rheumatoid arthritis, Parkinson's disease and dementia.

Recruitment

Participants will be recruited through general practices, selected to cover a broad range of regions and socio-economic conditions. Some 500-600 practices will be required with experience of working in research projects, and use of computerised prescription records (90% of practices currently do this). Recruitment will be managed through a centrally co-ordinated network of regional centres. All patients aged 45-69 registered with a selected practice will be sent a brochure explaining the aims of the project, a consent form and questionnaire (see below) and an invitation to participate signed by their GP. Those accepting will be asked to complete a questionnaire and invited to an interview with a research nurse.

Information collected

As described in box 1, participants will fill in a questionnaire on risk and lifestyle factors, be interviewed, undergo a physical assessment, donate a sample of blood for genetic analysis, complete a 7 day 'dietary diary' and agree to be 'followed up' at various points. Researchers will automatically be notified of any participants who die of heart disease, stroke or chronic obstructive airways disease and those diagnosed with various forms of cancer. People contracting conditions such as diabetes and neurological disorders, or suffering heart attacks will be identified by periodic follow up through GPs or via hospital records.

Box 1 Biobank baseline measures

Questionnaire – covering areas including socio-economic status, age, sex, habits and lifestyle, diet, reproductive history, family history, past health, disability/impairment, psychological status and early life factors.
Interview – with a research nurse covering details of medical and surgical history and current medication.
Physical assessment – standard methods to measure blood pressure, lung function, pulse rate, height and weight. Participants will receive feedback on the results, and their GPs will be informed of any abnormalities.
Blood samples – a small portion of each sample will be used for a full blood count and to measure vitamin C levels. The rest will be frozen and kept as a long-term source of DNA (for genetic analysis) and other material.
Dietary data –participants will also be given a 7-day diet diary to complete at home.
Follow up – NHS central registers will routinely notify researchers of participants' deaths (with details of cause of death) and of cancer registrations. GP and hospital records will also be used to follow up participants who contract certain conditions, or who are admitted to hospital. GP records can also be used to provide information on drugs prescribed during the course of the study. Participants will undergo some form of re-survey after 5 years, to update information on exposures (e.g. smoking, drinking, exercise) and state of health.
Source: Draft Protocol for Biobank UK, Feb 2002

Project funding and management

Initial funding for the project of £45M was announced by the MRC, Wellcome Trust (£20M each) and Department of Health (DoH, £5M) in April 2002. Each of the regional centres will be responsible for recruiting from multiple general practices. Exact numbers have yet to be decided, but if there were around 10 regional centres, each would have to handle 50 or so practices to get the required number of recruits. Arrangements for central co-ordination are still under consideration. Scientific management of the project will be the responsibility of a private company (Biobank UK) under the directorship of a chief executive who will be responsible to the funding bodies. An independent body will be set up to monitor and oversee Biobank's activities, although it has yet to be decided how this body will be constituted and what its remit will be. It is envisaged that it will be responsible for setting out broad guidance on who should be allowed access to information and under what circumstances.

What happens next?

Biobank is still in its infancy – many details of its design and management have yet to be finalised. The next 18 months or so will see a number of key developments:

- Recruitment of a chief executive and setting up the oversight body and scientific management group;
- Identifying the practices from which patients will be recruited and setting up the regional support centres;
- Training of staff, development of information systems, and setting up pilot projects with selected GPs.

Once these structures are in place, recruitment of participants may begin. This is currently scheduled to start by the end of 2003 and take around 5 years. Researchers will then apply to use the data in Biobank to test hypotheses about the contribution of genetic and environmental factors to common diseases.

Box 2 The Data Protection Act 1998

The Act sets down eight enforceable data protection principles that apply to personal data. These state that data must be: fairly and lawfully processed; processed for limited purposes; adequate, relevant and not excessive; accurate; not kept longer than necessary; processed in accordance with the data subject's rights; secure; and not transferred to countries without adequate protection. While it is intended that all data held in Biobank will be treated in accordance with these principles, the Act provides for exemptions from one or more of these principles, for example:

- for safeguarding national security;
- if requested by a Minister of the Crown;
- for the prevention or detection of crime;
- for apprehending or prosecuting an offender;
- for assessing or collecting tax or duty.

Technical issues

A number of technical aspects of the project's design have been questioned. These include:

- Some epidemiologists have suggested that even 500,000 people may prove to be too small a cohort to tease out the individual contributions of various factors involved in complex multi-factorial diseases.
- The age of the cohort and variations in the quality of medical record keeping means that the environmental and medical data collected are likely to be somewhat unreliable compared with the genetic information.
- There is also concern that the project may lead to an over-emphasis on genetic factors. This is because existing business models are more likely to commercialise insights based on genetics (e.g. selling drugs to 'genetically susceptible' people) than those based on environmental or lifestyle factors (e.g. healthier diets or exercise regimes).

Ethical issues

Underlying principles

Genetic information is covered by the provisions of the Data Protection Act 1998 (DPA). As outlined in box 2, this lays down eight basic principles relating to personal data and establishes certain circumstances where personal data may be exempt from one or more of these principles. The Wellcome Trust has stated that data in the Biobank will be treated in accordance with the DPA.

The Human Genetic Commission's (HGC) consultation paper (Whose hands on your genes?, November 2000) and subsequent report (Inside information, May 2002) identified four underlying principles relevant to personal genetic information:

- Privacy – a person should *"..not be obliged to disclose information about his or her genetic characteristics*".
- Consent – "*genetic information about a person should generally not be obtained, held or communicated without that person's free and informed consent*".
- Confidentiality – "*genetic information should generally be treated as being of a confidential nature*".
- Non-discrimination – "*No person shall be unfairly discriminated against on the basis of his or her genetic characteristics*".

These principles raise a number of ethical and social issues for Biobank, which are examined below.

Box 3 Provision of information

HGC considered what information should be given prior to consent being sought. It recommended this should include:

- the purpose and nature of the research including any physical procedures involved;
- storage arrangements for the information and how access to the database will be controlled;
- any implications associated with participation;
- likely future involvement of commercial interests.

The Wellcome Trust envisages sending information on the above to each potential participant, prior to inviting them to an interview where they will have the opportunity to discuss all aspects of the project. Anyone expressing misgivings about a particular aspect of the study will be advised to consider declining the invitation to participate.

Source: Inside information, HGC, May 2002.

Consent to what?

It is a fundamental principle of medical ethics that people receiving treatment or participating in research must first give their consent. In order to be legally valid, this must be freely given and fully informed. With respect to this latter requirement, it is not possible to predict all future applications for which Biobank might be useful. On the one hand Biobank research could be restricted to those applications that are foreseeable and for which detailed information can be given to people during the consent process. On the other, a more generalised form of consent may be acceptable, where people are given information on the type of research that may be conducted and asked to consent to their data being used for 'medical research'. HGC recently published guidelines on the extent of information needed for people to take informed decisions (box 3).

Biobank funders favour general consent, arguing that this will allow Biobank to be used to its maximum potential. They suggest there will be adequate ethical safeguards to protect participants: a research ethics committee will review all research proposals and each will have to comply with guidance laid down by the oversight body. However, groups such as GeneWatch UK, the Consumer's Association (CA) and Human Genetics Alert (HGA) question whether such general consent can be considered as 'fully informed'. Such groups would like to see people given more specific information so they could:

- specify the diseases their data could be used to study (HGA is particularly concerned that Biobank should not be used to study behavioural genetics);
- choose whether to consent to use of their data in research funded by commercial organisations;
- be given the option to be kept informed of when and where their genetic information is being used.

With specific consent, any use of data other than that explicitly outlined in the consent process would require re-consent to be sought. This would involve contacting people each time a new research proposal came up for consideration, a practice that HGC considered to be impractical and possibly also intrusive. HGC has stated that the evolving nature of research using genetic databases means that consent cannot be fully specific.

Secondary use of data

Secondary use of data refers to subsequent use for some novel purpose other than that for which consent was originally sought. As already noted, Biobank has sought to minimise such situations by seeking general consent to use data in medical research. However, at some point it is possible that researchers will apply to use Biobank data for purposes that fall outside the general description of medical research. It is not clear how such applications would be dealt with but it is likely that the proposed oversight body would have a role to play. For instance, it could consider all such proposals or issue guidance on research that can be supported. While the Lords' S&T Committee recommended a national body be set up to take such decisions, groups such as CA see re-consent for secondary use of genetic data as essential.

Security

Biobank data will be anonymised; it will be stored and routinely used in this format. The Wellcome Trust has stated that all data will be treated in accordance with the DPA and that none will be released in a form that allows individuals to be identified. But the anonymisation process has to be reversible to allow individuals to be identified for follow up purposes. This raises the potential for release of data. HGC has recommended that "*operators of all genetic research databases should be required to take rigorous steps to ensure that unauthorised access or disclosures are prevented*". Biobank intends to address this by storing files that allow individuals to be identified separate from the data, and by restricting access to them. Details of the security measures have not been finalised. HGA has suggested that those responsible for reversing the anonymisation should be independent of Biobank's owners and users to avoid any conflict of interest. The CA has also called for an independent agency to be set up to take responsibility for anonymisation and to control access to the Biobank data.

Access to data and genetic discrimination

Data held in Biobank are potentially of interest to a number of parties other than researchers. For instance:

- ***The police*** - the DPA (box 2) contains exemptions allowing the police access to personal data to prevent or detect crime or to apprehend or prosecute offenders. The police have already gained access to research samples by means of a search warrant.
- ***Insurers*** - use of genetic information by insurers is currently subject to a voluntary moratorium agreed with the Association of British Insurers (ABI) until November 2006. DNA genetic test results will not be used by ABI members except where the tests have been authorised by the government's Genetics and Insurance Committee.
- ***Employers*** - HGC has found no evidence that UK employers are using genetic data for recruitment or occupational health purposes. The Information Commissioner has issued a draft code of practice on the use of personal (including genetic) data for employment purposes and the Government is planning a wider review of policy in this area by 2005.

Controlling police access
Exemptions in the DPA (box 2) mean that operators of genetic research databases cannot guarantee personal information will not be divulged for purposes other than that detailed in the consent process. The Information Commissioner has thus called for 'ring-fencing' of research databases, or an explicit statement given to participants when consent is sought that police may have access to research data. HGC did not endorse this second option, considering that it would "*seriously discourage participation*", but it did recommend that "*consideration be given to legal means of preventing access to biomedical genetic databases by police and other law enforcement agencies*".

Genetic discrimination
Participants in the Wellcome Trust's consultations were concerned that employers and insurers may seek access to genetic data in the Biobank. HGC has noted that "*genetic research databases should not be used for any purpose other than such research*" and that this should be achieved by legislation if necessary. There are also wider concerns that Biobank research will lead to the development of an array of genetic tests, and that use of these could lead to genetic discrimination. Rather than modifying existing anti-discrimination legislation such as the Disability Discrimination Act 1995, HGC recommended the Government to consider separate legislation to prevent genetic discrimination in the UK. It saw this as part of a long-term policy review on the use of personal genetic information in insurance and employment. GeneWatch UK, CA and HGA have also called for legal safeguards to protect participants in projects such as Biobank. They are particularly concerned that the appropriate legal safeguards should be in place before Biobank starts recruiting people. However, Biobank recruitment is scheduled to start during 2003, while the HGC's recommendations for new legislation to prevent genetic discrimination are linked to a longer-term policy timetable (2005-06).

Ownership and benefits
Ownership of the data held in Biobank will remain with the three funding bodies. Insights arising from analyses of data may be developed into new products by commercial companies. This has led to concerns over the balance between public and commercial interests. However the question of who should benefit from the setting up of Biobank is a complex one. The project requires the altruistic involvement of large numbers of people who will not directly benefit from the results, so it is reasonable to expect some of the benefits to be directed towards the wider public good. On the other hand, companies taking the risks involved in developing new products will also expect some reward.

This raises the question of how best to share benefits between commercial companies and the wider public good. It is anticipated that Biobank will seek some form of payment for granting access, although many questions about the balance of interests have yet to be resolved. For instance, it is not clear what form the payment should take. Most interested parties agree that an up-front fee would be best, although a share of future intellectual property rights has also been suggested. It has also yet to be decided whether arrangements will differ for publicly-funded, commercial or overseas researchers, and how any money raised will be used.

Oversight of the project
Many of the uncertainties surrounding the issues discussed above – particularly those of consent, use of data and the balance between public and private interests - could be addressed by a 'trusted third party' constituted to oversee the project. People enrolling into Biobank could entrust such a body to make decisions on their behalf about such issues. Indeed, this approach was suggested by the Lords' S&T Committee as a way of making decisions on secondary use of genetic data. It has since been endorsed by the HGC, which recommended an independent (i.e. separate from the owners and users) body to oversee the database. While the draft Biobank protocol includes a proposal for an independent oversight body as part of the central management structure, the composition, remit and funding arrangements of this body have yet to be finalised. Factors such as the balance between 'lay' and 'expert' members, whether the body sets general guidelines or considers individual research proposals, and how it is funded will be critical in establishing its independence and perceived trustworthiness.

Overview
There is consensus that Biobank has the potential to greatly improve understanding of the factors behind many important diseases of adult life. However, the project raises a number of important issues that have received relatively little parliamentary scrutiny. While the S&T Committees in both Houses have considered the general issues raised by collections of genetic information, there has been only a short adjournment debate in Parliament (3rd July 2002) about the specific issues raised by Biobank. Wellcome Trust plans to continue its consultations on Biobank, with the results feeding back into the project's design. It recently announced an innovative new consultation asking people to make choices between different Biobank scenarios. However, HGA, CA, and GeneWatch UK would like to see consultation to address issues of consent, confidentiality and ownership before the project starts recruiting. Such groups do not oppose the project, but are concerned that it should not start until:

- Legal safeguards to prevent genetic discrimination are in place;
- A wide-scale public consultation has established an ethical framework for the project;
- Arrangements for independent oversight are finalised.

POST is an office of both Houses of Parliament, charged with providing independent and balanced analysis of public policy issues that have a basis in science and technology.

The Parliamentary Office of Science and Technology, 7 Millbank, London SW1P 3JA Tel 020 7219 2840

www.parliament.uk/post/home.htm