A new model of innovation in biomedicine?

A review of evidence relating to the changing relationship between the private and public sector in the use of human genomics and personal medical information

Paul Martin (University of Sheffield) and Greg Hollin (University of Nottingham)

November 2014

1. Introduction

This report was commissioned by the Nuffield Council on Bioethics to provide evidence for its inquiry into the Governance of Biological and Health Data with the following purpose - to assist the Working Party to understand:

1. The nature of the relationship between the public and private sectors in the development and execution of biological and health research
2. Whether there is an identifiable change in the nature of that relationship, e.g., from competition (Human Genome Project; BRCA identification) to collaboration (Genomics England)?
3. If such a change can be identified, what relationship, if any, does it have with the current conjunction of open data/open policy making?
4. Has the nature of public sector involvement in research changed?

Given the broad scope of this remit, the report has sought to chart changes in the relationship between the public and private sector by focussing on a series of key examples of developments in the field of large scale health data and human genomics. It will use these as case studies to consider the broader social, ethical and policy implications of any reconfiguration in the relation between the public and private sector.

The report is structured as follows. Firstly, it will start by placing any changes in the relationship between public research and private industry within biomedicine in a wider historical and policy context, including important changes in the production of knowledge and the rise of the knowledge economy (Section 2) and the development of biotechnology and genomics (Section 3). This will be followed by detailed analysis of three contemporary large-scale projects that use personal medical information and/ or genomic sequence data (UK Biobank; Genomics England; and care.data). These have been chosen as they are high-profile projects at the cutting edge of research and innovation that raise many issues. Section 5 will then consider the ethical, social and policy implications raised by the case studies in relation to the changing relationship between the public and private sector. Some final reflections on this shifting boundary will be offered in the conclusion (Section 6).

2. Models of innovation in biomedicine

Two major changes have occurred in the relationship between industry and academia in the post-War period and have accelerated in recent decades. Crucially both of these changes have been hypothesised to have had significant consequences for research and innovation. Firstly, there has been a discernable shift towards what has become known as a Knowledge Economy, or Knowledge Society. Government and, to some extent, industry have come to see academic outputs as being of not only intellectual but also significant economic value; science, technology, and their outputs being identified as solutions to, for example, the energy crisis of the 1970s and the economic crises of the 1980s and 1990s (Shinn 2002: 599). Simultaneously, universities have become aware of the economic value of the intellectual property produced within their walls (Nowotny et al. 2003: 182) and have sought to protect and make profit from those products of research.

Secondly, there has been a huge increase in the number of individuals passing through institutes of higher education. In 1950 less than five percent of young people in the United Kingdom enrolled in higher education (Boliver 2010: 232). By the first decade of the twenty-first century that figure had increased to forty-three percent, with numbers likely to increase still further (Chowdry et al. 2008: 2). This ‘massification’ of higher education (Gibbons et al. 1994: 70) has, it is claimed, lead to a significant number of graduates, surplus to the requirements of the academy, and yet capable of

producing highly profitable knowledge within industry and the public sector. The combined effect of these twin processes - the emergence of a knowledge economy and a large number of graduates able to contribute to it - has marked an important historical change.

This has been recognised at both national and international level, with a number of organisations advancing the idea of the knowledge economy in the 1990s and early 2000s, including the OECD, the European Union (Felt 2007) and the World Bank. For the latter this 'New Economy' is marked by increased codification of knowledge and development of new technologies; a high rate of science-based innovation; the importance of education and skills in the labour force, investment in intangibles (R&D, education); the significance of innovation for competitiveness; and greater globalisation and competition. In the UK during the 2000s a broad policy consensus coalesced around a commitment to these goals as a means of supporting the growth of the knowledge economy and has persisted to the present.

Several conceptual frameworks have sought to understand this changing landscape in which the lines demarcating industry, government, and the academy have been increasingly eroded (Hessels & van Lente 2008: 743). The notion of 'Mode 2 knowledge production', first described in The New Production of Knowledge (Gibbons et al. 1994) and the idea of the 'triple helix' of industry-government-academy relations (e.g. Etzkowitz & Leydesdorff 2000) have been the most influential of these. Michael Gibbons, Helga Nowotny, and the other authors associated with the concept of Mode 2 set-up a clear dichotomy between this novel form of knowledge production and that which went before it (Mode 1). Mode 1 is captured in the image of an Ivory Tower; esoteric questions of interest only to a specific disciplinary community were asked, answered, and evaluated inside those ancient spires (Gibbons et al. 1994: 3). By comparison, Mode 2 is associated with "(a) the 'steering' of research priorities [by society], (b) the commercialisation of research, and (c) the accountability of science [to society]" (Nowotny et al. 2003: 181). With Mode 2 the walls have fallen and the production, evaluation, and utilisation of science has been democratised, with science speaking to society, and society to science (Nowotny et al. 2003: 190). It also marks a much closer relationship between the public and private sector.

The notion of the 'triple helix' of industry-government-academy, whilst both distinct and empirically driven to a far great extent than the concept of Mode 2 (Hessels & van Lente 2008: 750), is similar in many respects. As the moniker's biological metaphor implies, the triple helix proposes an interweaving of industry, government, and academy relations; each influenced by the others and a totality emerging from this based on a 'nonlinear' model of interaction (Etzkowitz & Leydesdorff 2000: 114). In a continuation of the biological analogy, the authors see knowledge within the triple helix as evolutionary; there is no longer a frontier to knowledge, but an 'endless transition' whereby questions arise in response to the novel environments arising within a particular society (Etzkowitz & Leydesdorff 2000: 110).

The concept of Mode 2 in particular has not been immune from criticism with the historical, conceptual, and empirical validity of the core concept outlined by Gibbons et al called into question (see Hessels and van Lente (2008) for an overview). What is certain, however, is that both the notions of Mode 2 and the triple helix have become widely influential, within diverse circles, over the past twenty years. It is also true that, from the first articulation, health, biomedical, and biotechnical research have been identified as a particular site where Mode 2 knowledge production and triple helix forms of innovation can readily be found (Etzkowitz & Leydesdorff 2000: 120; Gibbons et al. 1994: 147; Shinn 2002: 605).

Many other theories of innovation have been developed in subsequent years and a full review of this literature is beyond the scope of this report. However, the point in mentioning these two examples is to illustrate the way in which new models of the relationship between the public and private sector within technological innovation have been important in shaping policy related to the knowledge

economy. The way this has played out in UK policy related to bioscience research and innovation will be discussed in the following section.

3. The changing relationship between the public and private sector in biotechnology and genomics

3.1 The birth of biotechnology and a new relationship between the public and private sector

The pharmaceutical industry has historically had strong ties with academia that date back to the early years of its creation in the late 19th and early 20th century. These ties grew in the post-war period with academic researchers providing fundamental knowledge that subsequently formed the basis for many highly successful medicines. Many of these relationships between universities and large national firms were long-standing and informal (e.g. ICI and University of Manchester), and this only started to change significantly following the birth of the biotechnology industry in the late 1970s.

The first wave of biotechnology firms created in the US were founded by leading academic scientists working in collaboration with venture capitalists and followed the model successfully adopted in the ICT sector in Silicon Valley. These firms established strong links with academic groups which were encouraged by an important legal change designed to stimulate innovation. The Bayh-Dole Act (1980) allowed intellectual property created using US Federal funding to be owned by universities and private firms as a means of encouraging technology transfer and commercial development (Mowery 2001). This had a major effect in providing powerful incentives for the commercial exploitation of academic research and was further enabled by the rapid development of a series of other initiatives designed to foster this goal. These included the creation of university technology transfer offices which were active in licensing academic research to both small new firms and established large pharmaceutical companies. In addition, public policy sought to encourage the growth of spin-out companies from universities as a means of valorising intellectual property created in the public sector.

The combination of major federal support for bioscience research, mainly through funding from the National Institutes of Health (NIH), and policies designed to support the creation of dedicated biotechnology firms (DBFs) lead to very rapid growth in the number of SMEs working in this field. A similar set of developments occurred in the UK and Europe, but lagged some five years behind the US. As this nascent industry developed in the late 1980s and early 90s, a new form of industrial organisation also started to evolve, with large companies increasingly forming collaborations with these new entrants (Malerba and Orsenigo 2002). At first, most of these alliances were concerned with licensing of technology or contract research and manufacturing, but by the late 1990s large drug companies were increasingly outsourcing core R&D activities. This led to their growing dependency on DBFs to provide early stage drug candidates that could be taken into human clinical studies (Pisano 2006). At the same time, biotechnology firms were also very active in establishing collaborations with academia. This was often initially with founding scientists, but increasingly collaborations were created with other leading research groups working in related fields. These took many forms including licensing, contract research, funding of students, sharing of equipment and resources, and consultancy. A significant effect of these developments was the increasingly tight linkage between academia and industry. Although this was common in other technological fields, such as ICT, new materials and engineering, the extent of these linkages within biotechnology was greater than almost anywhere else, raising a series of ethical issues (e.g. conflicts of interest and the (mis)direction of public research) (Owen-Smith & Powell 2011).

3.2 The early development of genomics: competition between the public and private sector

The first large scale human gene mapping and sequencing was undertaken in the late 1980s in the public and non-profit sector by the NIH in the US and Genethon in France. The commercial potential of this technology soon became apparent and led to a controversial set of patent applications in 1991 by Craig Venter, who was at the time a researcher at the NIH. These claimed ownership of partial gene sequences known as expressed sequence tags (ESTs) and were heavily criticised on a number of scientific and ethical grounds. Venter subsequently left the NIH to create one of the first dedicated genomics company (Human Genome Sciences) in 1992. This was followed by the establishment of a series of new genomics-based firms over the next few years that sought to commercially exploit the newly found ability to rapidly sequence large numbers of human genes, many of which it was hoped would form the basis for new drugs and diagnostics.

In 1990 the publicly funded Human Genome Project (HGP) was launched to sequence every gene in the human body with the aim of making major improvements in health and healthcare through improved diagnostics and the development of novel therapies. However, within a short time, a major concern was that the newly founded genomics industry would rapidly monopolise the use of human gene sequence data by claiming extensive patent rights. It was therefore hoped that by putting these sequences in the public domain it would become impossible to subsequently claim intellectual property rights on human genes. This in effect led to an international race between the public and the private sector to sequence and patent human genes on an industrial scale.

During this period there was massive investment in genomics by large pharmaceutical companies who sought to develop internal capacity and competences in what was increasingly perceived to be a key strategic technology (Hopkins et al 2007)). In particular, there was a widely held belief that genomics would provide a large number of new drug targets that would help address the ongoing crisis of productivity within the pharmaceutical industry (Ibid.). This investment involved both large numbers of collaborations with specialist genomics firms and universities, and building in-house gene sequence and identification facilities.

As the sector developed there was an important shift from sequencing to techniques that sought to determine human gene function and possible use as a target for drug therapy (target validation). Key to this was the ability to link gene sequence information to clinical data as a means of establishing the relationship between alteration in a specific gene and disease risk. This took the form of genome-wide association studies (GWAS) and was the motivation for the creation of a new type of research infrastructure known as a biobank, which combined clinical and sequence data. In a survey undertaken in 1999 (Martin and Kaye 1999) five small companies were already establishing large DNA sample collections that could be linked to personal health information. In addition, by the early 2000s large pharmaceutical companies were starting to routinely take DNA samples during clinical trials of new medicines to look at the genetic basis of adverse drug reactions, so called pharmacogenetics (ibid.). The most well-known proposal for the creation of a large national biobank was made by the Icelandic company deCODE Genetics which sought to combine personal health information, genealogical records and DNA sequence data on all Icelandic citizens. Although this project was never fully realised in practice, it established a model for how a private biobank might operate. In the late 1990s and early 2000s a series of other commercial biobanks were created at national and regional level. In 1999 proposals were made in the UK for the creation of a national biobank funded by the public sector to undertake epidemiological and genomic research; this subsequently became UK Biobank (see case study below).

The more general rationale for creating a large infrastructure of this sort was made in 1999 by Robin Fears and George Poste; the latter was at the time President of Research and Development at the

multinational pharmaceutical company SmithKline Beecham. They wrote an article in *Science* arguing that:

> "The United Kingdom National Health Service is a high-quality health care system, providing a comprehensive service to 59 million people, which also has the potential to serve as a unique resource for population genetics research. Capitalizing on the, as yet, untapped value will require appropriate scientific and clinical skills matched with large-scale computational infrastructure and proactive, transparent, and coherent policies for addressing the ethical, legal, social, and political issues arising from the use of clinical information. Public-private partnership is essential to maximize this opportunity to ensure that the clinical benefits of genetic epidemiology are realized and forge the evolution of increasingly rational care." (Fears and Poste1999, pp.267-268)

This vision of the NHS as a source of untapped value that could be realised through the creation of a large-scale computational infrastructure organised via a public-private partnership continued to guide policy for the next 15 years and formed the backdrop to the three case studies described in this report. It should also be noted that the first major controversy about commercial access to anonymised NHS medical records also took place in this period when a private company, Source Informatics, sought to analyse pooled patient information to help guide the marketing of prescription drugs. An initial ruling by the High Court found that such use did breach patient confidentiality, but this was overturned on appeal, allowing access to anonymised patient data to continue (Richards 2000, p. 77)

3.3 Industrial consolidation and the increasing importance of the public sector

The landscape surrounding the commercial development of genomics changed significantly following the bursting of the biotechnology 'stock bubble' in 2000/2001. The high valuations placed on genomics firms enabled them to raise large sums of capital in the late 1990s, but after 2001 the financial prospects for the industry seriously deteriorated. The situation became more difficult following the completion of the human genome project in 2003. Whilst this was a landmark in biomedical research that fully validated the significance of genomics as a field, it placed a large amount of gene sequence data into the public domain making further patent claims on human genes much more difficult. This undermined the valuation placed on companies that held large gene patent estates, as concerns grew about both the legal basis and commercial value of these assets. It also became clear that the main reason why many companies had filed human gene patent claims was to prevent competitors from restricting their research on particular drug targets rather than seeking to exploit the gene sequence itself .

Despite this, the genomics sector continued to grow until the mid-2000s, although the focus shifted to so-called functional genomics concerned with drug target identification and validation. By 2004 there were over 400 dedicated genomics companies working internationally. However, this number started to decline in subsequent years and the industry went through a phase of significant consolidation. A major reason for this was a change in the commercial strategies adopted by large pharmaceutical companies who started to experience a high rate of clinical failure in the first generation of new drugs derived from genomics (Bains 2004). This led to a reduction in genomics-based collaborations and some disinvestment. By the end of the 2000s the genomics sector had consolidated and was composed of a small number of first and second generation firms focused on drug product development, a larger number of SMEs working on functional genomics, and a series of other specialist companies developing next generation sequencing technology. Almost none of the commercial biobanks survived as it proved very difficult to find a sustainable business model, leaving these activities very largely based in the public sector.

One consequence of this was an increasing recognition that translating the promise of genomics into new drugs and diagnostics would take much longer that initially expected and that the work of the public sector would provide an important foundation for these activities.

3.4 UK policy to foster innovation and collaboration: building public-private partnerships

Throughout the 2000s a series of UK and EU public policy initiatives were taken to promote innovation and the growth of the knowledge economy in general, and the commercial development of biotechnology in particular.

In the fields of biomedicine and biotechnology high policy expectations surrounded the potential of genetics and genomics leading up to the sequencing of the HGP. In January 2002 the Department of Health announced the creation of six Genetics Knowledge Parks. These would act as centres of excellence that brought together industry, academics and clinicians to work on transferring genetic knowledge into health benefits for patients and the public health. This was one of the first attempts to improve genomic knowledge translation by changing the institutional, organisational and geographical relationship between the public and private sectors. This was followed in 2003 by the Government's White Paper on genetics in the NHS (Department of Health 2003) in which the Secretary of State for Health, John Reid stated:

> "Advances in human genetics will have a profound impact on healthcare. Over time we will see new ways of predicting and preventing ill health, more targeted and effective use of existing drugs and the development of new gene-based drugs and therapies that treat illness in novel ways. Above all, genetics holds out the promise of more personalised healthcare with prevention and treatment tailored according to a person's individual genetic profile".

The White Paper fostered a vision of the 'revolutionary' potential of genomics to transform healthcare and gave a strategic commitment to the development of genetics in the NHS, including the expansion and integration of genetic knowledge into mainstream services and clinical practice.

During the early 2000s an important shift also occurred in how the health service was viewed within policy. Historically, the NHS was primarily seen as serving the public health and facilitating economic development indirectly by ensuring a healthy population and fit workforce. However, with the emergence of the notion of the knowledge economy, the NHS started to be increasingly seen as a source of direct wealth creation. An important example of this shift was the 2003 report of the Bioscience Innovation and Growth Team (BIGT) 'Bioscience 2015 – Improving National Health, Increasing National Wealth'. The title of which clearly articulates the idea that these two goals of improving health and increasing wealth could be achieved within the same policy framework. The Executive Summary of the first chapter titled "Build a mutually advantageous collaboration between the NHS and industry for patient benefit" states that:

> "The NHS should be a major source of competitive advantage for the UK bioscience sector. As the main provider of healthcare to the UK population, it acts as a gateway to the largest aggregated patient pool in the world, and can monitor that population over time. This is a major asset for conducting efficient and effective clinical trials. The NHS is also a potentially significant source of innovation and early stage products for the commercial bioscience sector. … Ultimately, the BIGT wants to create a partnership between the NHS and the bioscience industry focused on a shared commitment to improved patient care through innovation." (BIGT 2003, Chapter 1)

This convergence of the interests of patients/ public health and the bioscience industry has remained a central motif with all subsequent policy.

One of the main recommendations of the report was improved co-ordination of clinical research and this led to the formation of the UK Clinical Research Collaboration (UKCRC) which brought together the NHS, industry, research funders, regulatory bodies, patient groups and academia. Another significant inquiry, the Healthcare Industries Task Force, was also initiated around the same time and reported in 2004 (HITF 2004). This sought to identify key ways in which greater collaboration between government and healthcare firms could be facilitated, particularly in the field of diagnostics. Specifically, it identified the way in which a comprehensive information technology (IT) infrastructure would be central to the development of what would later be called 'personalised medicine'. The Task force also echoed the BIGT report by arguing that the NHS was a uniquely valuable asset which could "provide an engine for industrial development based on the knowledge economy" (Ibid, Paragraph 4.35).

Other key actors that contributed to forging a new relationship between academia and industry included the Wellcome Trust, which has developed a coherent strategy over the last 15 years that has gone beyond the traditional role of a medical research charity. The foundation for this was laid with the establishment of the Sanger Centre to sequence the human genome and the associated Wellcome genome campus at Hinxton near Cambridge in 1992. In subsequent years the Trust has taken a series of initiatives to promote knowledge transfer, and provided a source of venture capital and other support for the creation of spin-out biotechnology companies.

A new paradigm for policy thus started to take shape during the first half of the 2000s in which government policy became directly guided by a series of principles; firstly, that genomics held great potential for both improving public health/ patient care and developing the knowledge economy. Secondly, that the NHS was a source of economic value and innovation, but this could only be unlocked by overcoming a series of technical, cultural and organisational barriers. Thirdly, that the role of government was to facilitate these developments by building a clinical research and IT infrastructure to support innovation, and re-engineering the institutional relationship between the public and private sector to promote knowledge translation. These became the main tenets of a policy consensus that has remained intact until the present. This is well illustrated in the Innovation, Health and Wealth report published by the NHS in December 2011:

> "This is why the Government's Plan for Growth, published in March 2011, announced that the NHS Chief Executive would review how the adoption and diffusion of innovations could be accelerated across the NHS. The NHS Chief Executive's Review and this report forms part of a wider UK strategy for Health Innovation and Life Sciences, alongside a detailed review of the Life Sciences industry led by the Office for Life Sciences. This wider strategy, led by the Prime Minister, sets out a comprehensive plan to transform the UK health innovation and life sciences sectors ..." (NHS 2011, p7)

Similar policy commitments can be clearly seen in other contemporary policy initiatives, such as the report of the Human Genomics Strategy Group in 2012 (Human Genomics Strategy Group 2012), which reiterates the transformative role of genomics and the need to create major new infrastructures within the public sector to enable this to happen. The next section describes three examples of large scale projects based on the creation of new data sets that have been driven by this broad policy vision and are actively collaborating with the private sector.

4. The contemporary relationship between the public and private sector in biomedical research

The way in which the relationship between the public and private sector is being configured in contemporary biomedical research will be examined in more detail using three detailed case studies: UK Biobank, Genomics England, and care.data. These have been chosen as they represent major high-profile initiatives, take different organisational forms and each aims to create a working relationship with the private sector. In particular, the case studies will explore how this is being achieved and the extent to which partnership goes beyond well-established forms of collaboration.

4.1 UK Biobank

UK Biobank (UKBB) is a large resource designed to support biomedical research by combining genomic sequence data with a wide range of other personal health information, with its primary aim being a resource for case control studies of common conditions. UKBB state that:

> "UK Biobank will help scientists better understand a wide range of chronic, painful and life-threatening diseases that typically strike from middle age. This includes illnesses like cancer, heart disease, stroke, diabetes, arthritis, osteoporosis, Alzheimers disease and other forms of dementia, oral and intestinal illnesses, eye disorders and depression." (UK Biobank 2014e)

Baseline data for inclusion in UKBB was collected between 2006 and 2010, with over 500,000 individuals aged between 40 and 69 recruited to take part in the study. However, it should be noted that the response rate was only ~10% of all those approached. Participants attended one of 22 purpose built assessment centres located throughout England, Wales, and Scotland. Once in the assessment centre information was obtained:

> "…on a participant's health and lifestyle, hearing and cognitive function, collected through a touchscreen questionnaire and brief verbal interview. A range of physical measurements were also performed, which included: blood pressure; arterial stiffness; eye measures…; body composition measures; hand-grip strength; ultrasound bone densitometry; spirometry; and an exercise/fitness test with ECG [electrocardiograph]. Samples of blood, urine, and saliva were also collected." (UK Biobank 2014c, p.2)

Not all of this data was collected for every participant as some additional measures (such as the arterial stiffness, saliva sample, and exercise test) were incorporated part way through the study. Similarly, several measures were abandoned as they were deemed to be too time intensive (a light-memory test, for example).

In addition to data collected at the assessment centres, information on participants is received from several other sources. Perhaps most prominently, participant data is linked with health records both pre- and post- assessment; death and cancer registrations are regularly obtained, as are Hospital Episode Data held by the Health and Social Care Information Centre (HSCIC)(UK Biobank 2014c, p.7). UKBB also occasionally request further information from participants (an online dietary questionnaire, for example, is currently in circulation and an imaging project has recently entered a piloting phase with the aim of ultimately collecting data from 100,000 UKBB participants) and, as participants were informed during their visit to the assessment centre, researchers using the database can request additional measures to be taken at later times (UK Biobank 2013).

Several aspects of this data collection procedure are worthy of consideration. Firstly, the individuals recruited for UKBB were healthy and, secondly, no particular areas of healthcare were identified of being of particular interest; the information obtained within the assessment centre was assumed, *a*

priori, to be of relevance to healthcare. Nonetheless, it is difficult to predict the range of research projects which will use the UKBB resource in the future. UKBB's study design has, at times, been subject to critique and described as being justified in a post-hoc fashion (Frank et al. 2006, p.596). It is also important to note that much of UKBB is, as described above, prospective in nature; relating to research that has yet to be imagined, data that has yet to be collected and, indeed, life events that are yet to occur. It is therefore essential that data collected by UKBB can be linked back to specific individuals. Data held by UKBB is, then, 'reversibly anonymised' (UK Biobank 2007a, p.11) meaning that UKBB data and participant information are held separately but can be linked via a securely held code. There are 'severe restrictions' placed upon the use of data which is not in anonymised form (UK Biobank 2007a. p.5) but it remains possible for participants to be identified. Furthermore, consent is 'broadly' given by participants to all research, by any party and at any time in the future, that "may, in the long term, help enhance other people's health" (UK Biobank 2007a, p.5). This prospective design and broad consent raise particular ethical issues.

Organisation of UK Biobank

UK Biobank is a not-for-profit charity (in England, Wales, and Scotland) based in Stockport and managed by the University of Manchester. UKBB was given the official go-ahead in 2002, although its origins lay in a series of policies aimed at supporting the development of the biotechnology sector in the late 1990s. It was initially launched with £61.5 million funding given by a range of charitable (Wellcome Trust (WT), British Heart Foundation (BHF), Diabetes UK) and governmental sources (Medical Research Council (MRC), Department of Health (DoH), Scottish Government, and the Northwest Regional Development Agency). This initial funding base has been extended in the decade since the formation of UKBB; an additional £6 million was allocated for further data collection and £25 million has been allocated for use between 2011-2016 to assist with the creation of online access and data storage facilities (UK Biobank 2014a). Recently (in March, 2013) an extra £21 million (from the DoH £10 million, MRC £10 million, and BHF £1 million) has been awarded to allow genotyping to take place on the blood samples already stored by UKBB. This genotyping is being undertaken by a private firm, Affymetrix[1] and is anticipated to be available to researchers in 2015. Furthermore, the MRC has recently awarded £9.6 million to conduct new magnetic resonance imaging of the brain, heart and abdomen, low power X-ray imaging of bones and joints and ultrasound of neck arteries from 100,000 UKBB participants starting in 2014 (UK Biobank 2014d). This brought the total of committed funding to the project to in excess of £110 million, making it the second biggest single biomedical research project in UK history after the sequencing of the human genome. Currently, there is no attempt being made to recoup any of these start-up costs from users of the UKBB (UK Biobank 2014f).

In terms of governance "A board of directors, accountable to the members of the company (the Medical Research Council and The Wellcome Trust), act as company directors and as charity trustees. They have overall responsibility for the direction, management and control of UK Biobank…" (UK Biobank 2014h). This ensures effective accountability to the main funders of the project. The Board is largely made up of academics and policy makers with one representative from industry.

How will public and private sectors collaborate within UKBB?

The model of public-private partnership (PPP) employed within UKBB is reasonably straightforward. All data is collected and stored by UKBB and housed in Stockport. Any organisation which wishes access to UKBB data pays £250 + VAT as an administration charge and a

[1] This appears to be the first time that UKBB has commissioned a private company to analyze data, something which is very much at the core of Genomics England (GeL).

further £1,500 + VAT per successful application which requires only access to the dataset. Any request which requires the use of depletable biological samples (i.e. blood, saliva, or urine) or which requires in-situ access incurs a further cost, determined on a case-by-case basis. These costs are the same regardless as to whether the bid comes from the charitable, academic, or private sector, the UK or overseas[2]. Further, UKBB will not seek to profit from any research which emerges following use of the data sets, or prevent patenting by private companies[3]. In short, therefore, UKBB does not so much *collaborate* with industry[4] as provide *non-discriminatory*[5] *facilitation* to the 'unique resource' that is the NHS patient pool and the data that BBUK has gathered.

While UKBB expect biotechnology and pharmaceutical industries to be significant users of the resource (UK Biobank 2007a, p.18), it should be noted that since the 30th of March, 2012, when UKBB was made available to researchers (UK Biobank 2014b), just over forty research projects have been approved[6] and all these projects are based within academia.

One recent development that should be noted is that the team behind UKBB have started a spin-out company, offering a range of services, known as the UK Biocentre[7] (UKBC). The services offered by UKBC fall, primarily, into two categories. Firstly, UKBC state that knowledge obtained during the foundation of UKBB (concerning, for example, ethics, participant recruitment, and logistics) can be put to use by others conducting similar projects (UKBC are playing a role in the creation of a Saudi Biobank, for example[8]). Secondly, UKBC offer facilities for data storage and analysis, with research projects storing their samples in Stockport. However, Biocentre South, based in Oxford, will not be open to use by the private sector. The INTERVAL study, a joint endeavour between the Universities of Oxford and Cambridge, as well as the NHS, is an example of a project utilising UKBC for just such data storage and processing (UK Biocentre 2013). It is claimed that any profits made by UKBC will be put back into UKBB in order to strengthen the resource (UK Biocentre 2014).

Social and ethical issues

In order to address the many social and ethical issues raised by UKBB a comprehensive Ethics Governance Framework (EGF) has been adopted, with an independent Ethics and Governance Council established to "advise the Board and Funders, and publish public reports on the conformance of UK Biobank with this Ethics and Governance Framework and with the interests of participants and the public." (UK Biobank 2007a, p.5). Membership of the Council draw from a range of expertise in science, law and ethics. The main issues covered by the EGF, and that are of relevance to this report, concern consent, anonymity and data sharing. Given that UKBB states that it is "impossible to anticipate all future resource uses" (UK Biobank 2007a) this raises concerns about how participants can ever be truly give informed consent. UKBB attempt to address this issue of informed consent by, firstly, making participants aware that UKBB is a research resource and not a healthcare programme. Secondly, at the time of data collection, the EGF states that participants should be made aware that the information obtained by UKBB would be linked-up with data from elsewhere in the NHS. In terms of data sharing and access, as the recent care.data project has demonstrated, the types of data that are held centrally

[2] It was initially stated that the commercial sector may have to pay more to access the data sets (UK Biobank Interim Advisory Group on Ethics and Governance 2003, p.15; UK Biobank 2007b, p.98-99) but that idea appears to have been abandoned by the time that researchers were given access to the data in 2012.
[3] For two declared reasons: 1) Research generally relies on numerous sources of input so would be difficult to claim; 2) Private enterprise is deemed less likely to use the resource if profits are being shared (UK Biobank 2014f).
[4] With the exception of the recently commissioned work provided by Affymetrix.
[5] UKBB do not, for example, exclude the tobacco industry for access (UK Biobank 2014g).
[6] https://www.ukbiobank.ac.uk/approved-research-2/
[7] http://www.ukbiocentre.com/
[8] http://kaimrc.med.sa/index.php?option=com_content&view=article&id=31&lang=en

within the NHS, and are thus able to be linked together, is continually changing. Participants are therefore informed that the nature of the data to be linked would be determined at a later point. However, data linkage will be governed by the EGF which is a statement of ethical principles and governance which acts as a long-term commitment to both participants and the public. At present, UKBB is already accessing Healthcare Evaluation Data (HED)[9] held by the Health and Social Care Information Centre (HSCIC). Furthermore, given that the HSCIC are in the process of combining HED with GP records, it is reasonable to assume that UKBB will be granted access to at least some portions of the integrated data set being created by care.data (see below). However, at present there is no indication that UKBB is forging links with Genomics England.

The issue of the nature of the PPP has also been of concern to those involved with UKBB's EGF. As early as 2000 it was noted that many members of the public were 'alarmed' at the prospect of industry accessing and then profiting from UKBB (Dawson 2000, p.9). A public consultation on UKBB's access procedures conducted in 2011 found that commercial use of UKBB was still a particular concern, albeit lagging someway behind issues surrounding security, re-contact, communication, and assessment feedback (UK Biobank 2011b). In response to the suggestion that academic researchers, or those within the UK, might have preferential access to the resource UKBB stated that "it has been made clear that the Resource is to be made available to all bona fine researchers without preferential or exclusive access, including those working outside the UK or for commercial organisations," and that this was made explicit in the information and consent materials provided to participants (UK Biobank 2011b, p.7).

4.2 Genomics England

Genomics England Limited (GeL) is a government owned, for profit company, established in 2013 to run the 100k Genome Project. This project involves conducting 100,000 whole genome sequences (WGS) on patients recruited through the NHS in England. The 100k Genome Project, and the work of Genomics England, are surrounded by high expectations with the intention of it becoming a world leader in the arena of genomic medicine. The Executive Chair, Sir John Chisholm, described the project as "the most important medical step in the 21st century" and "in the order of the Human Genome Project" (Genomics England 2013g). In particular, it is anticipated that the 100k Genome Project will "deliver benefit to the community at large, both in terms of health and future wealth, and at the same time feeding back appropriate insights to the clinicians treating participating patients.[10]" (Genomics England 2013h). It is also claimed that the project will enable the UK to become the first country in the world to introduce th technology into its mainstream health system (Genomics England 2013a). The economic benefits are anticipated to emerge from several sources. Firstly, by "going straight to the answer rather than spending 20 years wandering around the hospital wards." (Genomics England 2013f) it is believed that the NHS will save money previously spent on erroneous diagnostics and treatments. Secondly, the pharmaceutical industry will come to Britain to use the project infrastructure and this will lead to the development of new diagnostics and treatments (Genomics England 2013e).

In terms of its scientific programme there are three specific areas of focus;

[9] Hospital Evaluation Data is an analytics suite which integrates Hospital Episode Statistics with the Office of National Statistics mortality data sets and both in- and out-patient data with the aim of allowing users to "monitor, compare and evaluate hospital performance indicators with NHS-wide connectivity." (Anon 2013, p.2)

[10] With regards to this final point, GeL is evidently Epsteinian (Epstein 1995; Epstein 1996) – patients who participate might expect *personal* benefit/insight, have the capacity to shape the project, and (following a question at the town hall meeting) it is explicitly stated that participants will be able to have a copy of their own data.

1) Cancers. WGS will be conducted for both the patients (searching for inherited markers) and also upon the tumours themselves (searching for de novo mutations). A decision has not yet been made on which forms of cancer to examine, although cancers of childhood may be targeted as these are less likely to be caused by de novo mutation and more likely to have an inherited aspect (Genomics England 2013g).

2) Rare conditions (defined as affecting <1:1500 people). These are likely to be Medelian conditions, although no decisions have been made as to which rare conditions will be included in the project.

3) Infectious disease with a focus upon HIV, tuberculosis, and hepatitis C. A particular area of concern here is antibiotic resistance.

In terms of the rationale for these initial targets the links between these three areas are: a) they place significant financial burdens on the NHS; b) there are realistic hopes for the capacity of genomic medicine to significantly improve understanding and treatment in these areas.

Participants taking part in the pilot stage of the 100k Genome Project sign different consent forms to those taking part in the main stage of the project. One of the most significant differences in that, during the pilot phase, data is distributed to commercial organisations and annotation of the whole genome sequences is conducted within the annotators' infrastructure, outside of the much vaunted NHS firewall (Genomics England 2013c).

It is widely recognised that individuals can be identified from large data sets like that which will result from the 100k Genome Project and, indeed, that possibility was realised within a research setting as long ago as 2008 (Narayanan & Shmatikov 2008). There is, of course, particular concern when the data set in question is comprised of whole genome scans given that "the human genome is effectively a unique identifier" (Wright et al. 2011, p.86). GeL have a dedicated framework in place in an attempt to ensure that participants cannot be individually identified by users of the 100k Genome Project. Firstly, data will be fully protected behind an NHS firewall (Palin 2013) which no data will cross[11]. At the conclusion of the data collection phase of the project academics, clinicians, and industry will be allowed inside the firewall in order to conduct analyses. However, these analyses will be conducted with pseudonymous data comprising of anonymous clinical data and DNA sequences and not with the readily identifiable data stored by GeL[12]. Further, at the conclusion of their work, users of the GeL dataset will only be able to take their results (and not the raw data) outside of the NHS firewall (Genomics England 2013g). During the project, data will be collected inside the firewall by sequencing organisations contracted to GeL (Genomics England 2013f).

Organisation of Genomics England

Genomics England (GeL) was formally launched by Secretary of State for Health, Jeremy Hunt, in July 2013. The UK government has pledged £100million in support for the project and there are expectations that philanthropic investors will also provide considerable funding; indeed, additional funding is a prerequisite as £100m will not cover the expected costs. The project is due to conclude in 2017.

[11] With the exception, as noted above, of those consenting individuals taking part in pilot phase of the programme.
[12] This data is pseudonymous because it is potentially, rather than immediately, identifiable. See Narayanan and Shmatikov (2008).

Legally GeL is unusual for a project of this sort in being incorporated as a private company entirely owned by the Department of Health[13]. The consequences of this are made clear in its web pages on 'How we will work':

> "The company's Business Plan will include a financial model which will describe how the project objectives are to be met and what resources will be required to achieve them. It will make an assessment of the trajectory of costs and make projections for the development of income.
>
> The company has some key parameters for how it will work. It will:
>
> - Only employ a small number of technical, procurement and business experts.
> - Quickly be established and get up to speed.
> - Be flexible, agile, able to move quickly as the market changes.
> - Participate in the market as a business talking to businesses.
> - Ensure the benefits of the investment flows from the company to a large range of companies and contractors including SMEs.
> - Use any surplus to benefit the public health community.
> - Have the Secretary of State as sole shareholder." (Genomics England 2014b)

What is notable here is the commitment to a commercial model of governance, planning and organisational behaviour driven by the market and benefitting industry (as well as public health). However, it should also be noted that as of April 2014 the specific mandate of Genomics England had still to be clarified, as it was not formally a research project, but neither was it involved in delivering clinical care or providing diagnostic services to patients.[14]

The company itself is primarily responsible for project delivery and data protection during the 100k Genome Project, but will contract out sample collection, analysis, and storage to UK based companies, universities, and hospitals (see below).

GeL state that it is intended that the data from the 100k Genome Project will be linked with identifiable data from primary care and hospital records (Lewis 2013); see section of this report on care.data. To achieve this GeL will work with partners (such as the Health and Social Care Information Centre, Public Health England, National Institute for Health Research and the research councils) to ensure that genomic information can be linked to relevant clinical records. This will mainly occur initially with major cancer centres to support a better understanding of clinical impacts, genetic variation and to measure outcomes more precisely (Keogh & Kelsey 2013, p.3). There do not appear to be any moves at present to connect GeL with UKBB, although the ethics advisory group are keen to learn from the (ethical) experience of both 'the spine'[15] and UKBB. In the future a parallel programme examining epigenetics may be launched in which participants' information might be placed into biobanks (Genomics England 2013f).

---

[13] Despite being a company, GeL is repeatedly described as being "by the NHS, for the NHS". In the first town hall engagement meeting John Chisholm states that it was decided that GeL should be a company "because of the difficulties and the need to get focus on it [the project]" (Genomics England 2013g). Later in the same meeting Mark Caulfield, the project's chief scientist, states that "Although this [GeL] is a company, it is only formed as a company so it can move more quickly to do these things [help patients, the NHS], to bring maximum benefit at the fastest speed" (Genomics England 2013g).

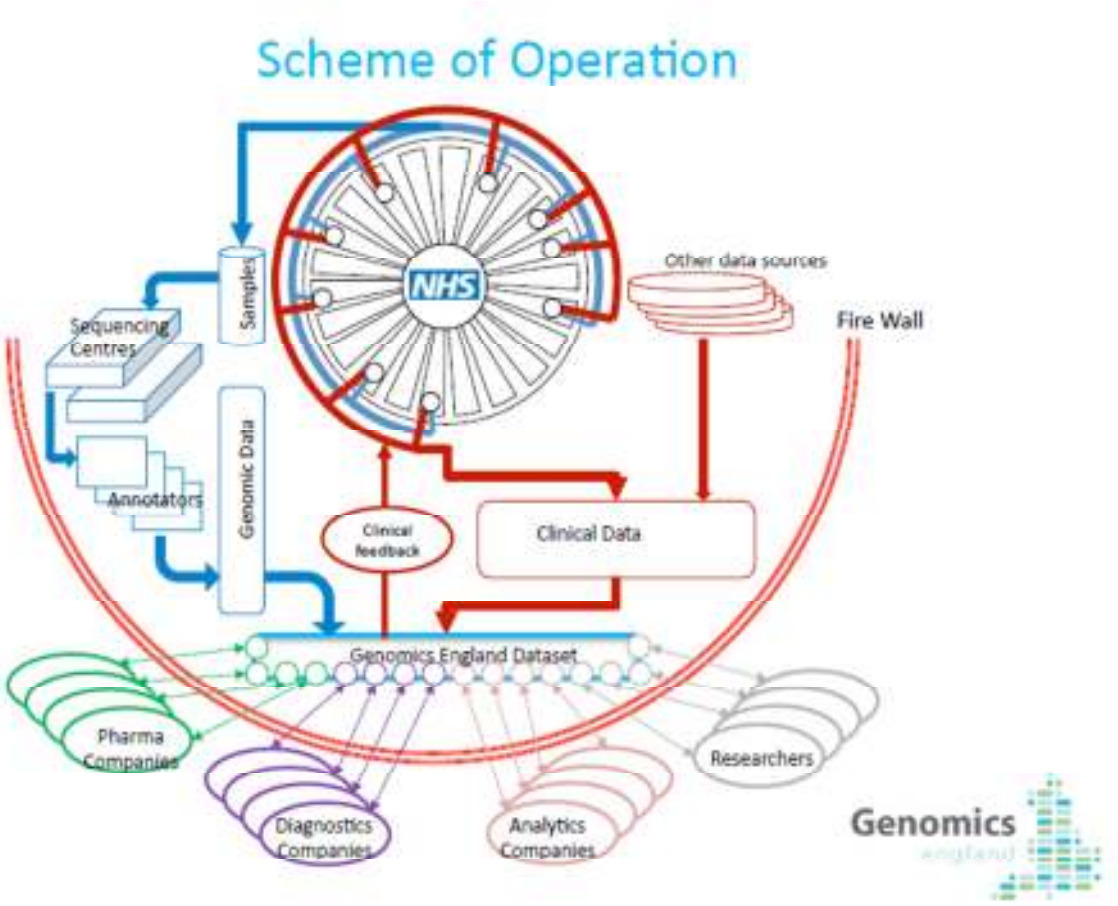
[14] Note: GeL state that "the core purpose of the project is to deliver benefit to the community at large in terms of heath and future wealth, and at the same time feeding back appropriate insights to the clinicians treating participating patients." Further, It was also specifically said in the annotators meeting that annotators will need to be quick with the turnout because we'll be dealing with 'real cancer patients' and their data will be needed

[15] The spine was a proposed central database of electronic NHS data which was abandoned after cost estimated to be in excess of £12billion.

The relationship between the public and private sectors

There are two distinct types of private sector companies which are integral to GeL; firstly, software developers and those focusing upon how to obtain meaning from the data (sequencers and annotators) and, secondly, those within the life sciences, such as pharmaceutical companies and those concerned with developing products. Both types of private company are well represented on GeL's Board. John Chisholm (Executive Chair) has founded a software company (CAP Scientific) and is also chair of the Medical Research Council); Kevin J. Dean (Non-executive Director) has a background in the innovation sector of the life sciences; and Jon Symonds (Non-executive Director) is on the Executive Committee of Novartis and has held prominent roles in AstraZeneca.
As stated above, for the duration of the 100k Genome Project GeL will contract out sample collection, analysis, and storage to UK based companies, universities and hospitals (although companies from both mainland Europe and the USA were present at the annotation supply event). The tendering process was already underway by March 2014 and software innovators have been invited to bid for £10 million of funding available to "analyse and better interpret significant variations in the DNA of NHS patients with serious illnesses" (Genomics England 2013i). The first contract, to sequence the data of the 2,000 patients involved with rare and inherited diseases pilot has been awarded to Illumina Sequencing (Business Wire 2013), although it is anticipated that other awards (including the sequencing of the cancer pilot of 6,000 individuals) will be awarded to multiple companies.

Companies focusing upon annotation have been informed that they may collaborate with GeL in numerous ways (Genomics England 2013b). Firstly, GeL are encouraging private companies to aid them in their commissioning process, helping to determine which companies are backed by GeL. Secondly, GeL are tendering for software suppliers concerned with both genome sequencing and annotation. Finally, there will be an "opportunity that you'll have is to participate beyond Genomics England in the process of integrating the learning from our data set into the clinical community" (Genomics England 2013b). This final point refers to those companies interested in accessing the final data set and, as the diagram below shows, is expected to include pharmaceutical, diagnostic, and annotation companies, as well as academic researchers.



(Genomics England 2013d)

As a consequence, there is an expectation of private/public collaboration both *during* the project (in the act of data storage and collection, for example) and *after* the project (as academics, clinicians, training bodies, and industry use the data made available). A key assumption underpinning the whole project is that the results of the 100k Genome Project and other GeL activities will be commercialisable. GeL anticipate profit to arise from the dataset in several ways: i) Companies may simply pay to access GeL's database of sequences; ii) GeL as a company, may enter into a royalty sharing scheme or pursue a joint venture model, options designed to aid small and medium sized businesses ; iii) companies may purchase exclusive time limited access to the dataset (Genomics England 2013d, p.8). These details are, at this time, still largely speculative.

Social and ethical issues

The organisation of the project and the close links with industry within GeL raise a number of important social and ethical issues. These are complicated by the uncertainty surrounding GEL's formal mandate, as it currently sits in something of a governance vacuum, being neither a research project not a provider of clinical care or diagnostics. In response, an Ethics Advisory Group (EAG) has been established and has produced an initial memo (Parker 2013a). Prof. Michael Parker (Chair, EAG) notes several areas of particular ethical consideration (Parker 2013a). The first issue is that of consent. It is explicitly stated in the EAG's memo that UKBB is a good example to follow and that consent should be 'broad', allowing access for a range of purposes from a range of bodies. Accordingly, participants should be made aware of potential problems concerning anonymity (see above) and that a range of different bodies (from the public, private, and charitable sectors) may access their data. A second ethical concern is that of data access. As with the issue of consent, there is concern here about the need to make participants aware that commercial organisations may utilise

the data set and that, in theory at least, this data may be identifiable. It is also noted that thought must be given to issues which may arise if participants wish to access to their own data. Finally, there are the related issue of feedback and public trust, confidence, and involvement in the project. These issues all relate to the need for effective oversight and governance and much of the EAG's memo is concerned with the form that such governance might take. As the project is still at an early stage, little appears to have been decided upon regarding these points, with Michael Parker suggesting that 'no ethical issues are off the table' (Parker 2013b).

Cutting across several of these strands, the advisory group note that public weariness over private/public partnerships is something that must be addressed (Parker 2013a, p.3). The key concern here – trust of patients, doctors, etc – in private/public relations and the message in the strapline "health and wealth" are well established. There is also scepticism regarding GeL's motives. As reported by Alistair Kent from *Genomics Alliance UK* during the town hall meeting, there is "hype [from some sections of the media] that this is a sinister conspiracy to get a DNA database for the whole population for nefarious purposes of government, or to support the uncontrolled application of industrial might into the generation of pharmaceutical industry profit" (Genomics England 2013g). It is to ward against these accusations that the board of GeL insist that the project is 'by the NHS, for the NHS' (Genomics England 2013g).

In response, advocates of the project, including the EAG, stress that private/public collaboration is central to GeL's mission and that maximising benefit to patients will require the involvement of public/private partnerships. Accordingly, commercialisation is deemed to be in the public interest. There is an awareness on the part of the EAG that a high level of trust will be required to ensure public confidence in the project and that patients, clinicians, and scientists are not deterred from taking part. With this in mind the project will have a robust governance framework, led by the EAG, and which is still under development. GeL, in collaboration with the Sanger Institute, have just launched an online 'Ethics and Genomics Survey', described as a first step in a thorough ethical analysis of the 100k Genome Project (Genomics England 2014a). Based around a series of videos, the survey is concerned with a number of areas relating to the sharing with participants of raw data, pertinent and incidental findings, as well as how such findings might be presented. There is also a consideration of whether researchers have a duty to search for such findings and some consideration of the 'flexible' consent that participants might be required to give.

4.3 Care.data

Care.data is a project commissioned by the Health and Social Care Information Centre (HSCIC), an executive non-departmental body within the NHS. The project relates to NHS England only, and different arrangements are in place for Wales and Scotland. Taking its current form following the Health and Social Care Act of 2012, HSCIC was created with the intention of being a national focal point for information collections across health and social care (Department of Health 2012, p.2) that is:

> "…responsible for collecting, transporting, storing, analysing and disseminating the nation's health and social care data. HSCIC is responsible for providing a trusted, safe haven for some of an individual's most sensitive data." (Health and Social Care Information Centre 2014a)

HSCIC aim to 'revolutionise' the ability to unlock NHS and healthcare data (Health and Social Care Information Centre 2012) and has been given the legal and administrative power and responsibility to:

i) Collect information from health and social care bodies;

ii) Hold that information within a secure environment;
iii) Make that information readily available for others to turn into "actionable business intelligence" (Health and Social Care Information Centre 2012)

This final point refers to the 'pan-government' commitment to make data available openly thus ensuring that 'patients, academics, data intermediaries and other organisations can access the data and make comparisons that enable more informed choices' (Health and Social Care Information Centre 2014e)[16]. The intention is that care.data will provide comprehensive and integrated information about the care patients receive from all parts of the health service, including hospitals and GP practices (Health and Social Care Information Centre & NHS England 2013). It is hoped that by pulling this information together and then analysing it in de-identified fashion, researchers can compare the safety of different NHS hospitals, monitor trends in different diseases and treatments, and use the data to plan new health services" (Lewis 2014). While HSCIC will not exclude particular organisations from using their data (e.g. insurance companies), access to data sets will only be granted if it is shown that the data will be used to benefit the health and social care system (Royal College of General Practitioners et al. 2014).

At present, HSCIC collect (monthly) Hospital Episodes Statistics (HES) which relate to in- and out-patient appointments, and accident and emergency admissions (Health and Social Care Information Centre 2014c). In order to create the new data set behind care.data a General Practice Extraction Service (GPES) will commence[17] in the Autumn of 2014 and, as with HES, continue hereafter on a monthly basis. Under the GPES scheme, information stored electronically within GP practices will be extracted and sent to the HSCIC (NHS England 2014). This information includes easily identifiable 'red data' (such as an individual's NHS number, date of birth, full post code) as well as a patient's history of diagnoses, prescriptions, vaccinations, and so forth. At present, particularly sensitive information (e.g. abortion history) and hand written notes will not be uploaded (Bhatia 2014). It is this integration of the hospital (HES) data set and the GP data set (from GPES) which constitutes care.data. In addition, it is anticipated that care.data will join up with other NHS projects, for example, allowing phenotypic data to be linked to the genomic data produced by the 100k Genome Project being co-ordinated by Genomics England (GeL) (Lewis 2013, p.5). Indeed, Sir John Chisholm, Executive Chair of GeL, has recently become a Non-Executive Director at HSCIC (Health and Social Care Information Centre 2014d). GeL state that it is intended that the findings of the 100k Genome Project will be linked with identifiable data from primary care and hospital records and that this can be linked to relevant clinical data (Keogh & Kelsey 2013, p.3).

Public and private sectors collaboration within care.data

The model of public-private partnership (PPP) employed by the HSCIC is almost identical to that employed by UKBB and is reasonably simple. HSCIC does not sell patients' data (Health and Social Care Information Centre & NHS England 2014, p.6), stating that it is:

> "…publicly funded and we therefore operate on a cost recovery basis. We do not charge for data itself but we do apply charges to cover the costs of processing and delivering our service. Where possible, HSCIC data is made available as open data." (Health and Social Care Information Centre 2013a, p.1)

HSCIC make data, including potentially identifiable data, available to various 'key partners', including industry, at a cost determined by the amount and type of data required (between a few

[16] See also the closely allied Clinical Research Practice Datalink.
[17] This will be undertaken by software company ATOS and, as with Affymetrix's genotyping in UKBB, this appears to be the only time a private company are involved in data collection.

hundred and a few thousand pounds (Health and Social Care Information Centre 2013a)). The only form of open collaboration[18] is the collection of the GPES data, which is done by the company ATOS before being securely stored by HSCIC.

HSCIC do not specify the type of researcher that they believe to be interested in their resources. The Data Access Advisory Group (DAAG), an independent body hosted by HSCIC do, however, make available details of all approved projects that utilised HSCIC's data sets with sensitive data. Sensitive data may include a patient's NHS number, postcode, date of birth and/or death, physical and mental health, and so on (Health and Social Care Information Centre 2013b). Since HSCIC's formation in its current guise in 2013, 68 requests for sensitive data have been approved and, while the majority of these requests have come from within the NHS or academia, a number have come from industry (Health and Social Care Information Centre 2014b). Civil Eyes Research Limited, a company specialising in metrics of quality and productivity in health services[19], have had requests approved in each of the last three years and CHKS and Dr Foster Intelligence, who have similar goals, have also had requests approved. Indeed, the research undertaken by Dr Foster Intelligence, utilising freely available data provided by HSCIC, which seems to demonstrate that mortality rates for elective surgery are higher at the weekend (Aylin et al. 2013), is widely regarded as one of the success stories of the open data movement. Other organisations directly involved with healthcare which have requests for sensitive data approved include, Corin Ltd (who specialise in innovation around orthopaedics), private healthcare provider BUPA, and UK Biobank. HSpot Ltd (a software company) and the Institute for Fiscal Studies have also had access approved. It is noticeable that the pharmaceutical industry is absent from the above list. However, and as noted above, GeL are seeking to link care.data with the 100k Genome Project which is explicitly targeting input from the pharmaceutical industry (Genomics England 2013d, p.7) and may thus be expected to join the bioinformatics sector as significant users of HSCIC data.

Social and ethical issues

A large number of social and ethical concerns have been raised about care.data. Some of those are broadly similar to those noted above in relation to UKBB and the 100k Genome Project. As with these projects, care.data employs a 'broad' or 'flexible' approach to consent whereby participants cannot opt out of particular types of data usage: it is an in-or-out model. Furthermore, much of the data contained within care.data are potentially identifiable (and include readily identifiable data). These two issues of flexibility and anonymity have been identified as ethical concerns for care.data (e.g. Bhatia 2014), although, taken in isolation, the issues raised are not noticeably different to those raised by the projects covered above.

However, where care.data does stand in stark contrast to both UKBB and GeL, is that it operates an opt-out model of presumed consent whereby individuals must inform their GP if they do not wish to take part in the scheme. This opt-out arrangement has been an area of particular criticism for a number of reasons. Firstly, many commentators and three quarters of GPs surveyed (Matthew-King 2014a) believe that active participation (an opt-in model) should be used. In response, proponents of the project stress that anonymised data is widely used in health services research without patient consent. The issue at stake here is therefore the extent to which this principle should be applied where data can be identifiable i.e. is 'pseudonymous'. The controversy around this led to a decision by the Secretary of State for Health, Jeremy Hunt, to provide an opt-out mechanism.

[18] Technically, GeL is also a company and, as discussed above, there is the possibility for both data sharing and profit here. However GeL does position itself as being 'within the NHS'.
[19] http://www.civil-eyes.com/

Secondly, there has been controversy about this opt-out process. The government ran a public information campaign via a posted leaflet entitled 'Better Information Means Better Care' to make people living in England and Wales aware of care.data and their right to opt-out. This campaign and the leaflet have been criticised as the term 'care.data' is not used in the publication and there is little detail of the project. Furthermore, an opt-out form was not included with the leaflet, with those receiving the leaflet invited to individually contact their GP if they do not wish their data to be included in the project. Finally, a poll conducted for the BBC suggested that only 29 percent of people recall receiving the leaflet (Triggle 2014). Furthermore, a recent Freedom of Information request showed that households that have opted not to receive junk mail would not have received 'Better Information'. Indeed, the BBC has reported that 45 percent of people were unaware of any of the activities undertaken in the GPES. Growing public criticism of care.data by both patient and professional groups led to a decision in February to delay its launch by six months until Autumn 2014.

Thirdly, a number of issues about the security of data have come to light following an NHS risk analysis which concluded that the data held could be vulnerable to hackers or could be used to identify patients "maliciously" (Donnelley 2014a). These concerns were compounded when a Freedom of Information request revealed a series of security breaches have occurred on centrally held records, despite public assurances by NHS senior executives that this has never occurred (Donnelley 2014b). These included the theft of an unencrypted laptop, holding details from the hospital records of more than eight million NHS patients, which could identify individuals.

Finally, in April 2014 it was revealed that HES data on every hospital patient had previously been sold to a major society of insurance actuaries and that this has been used to help insurers refine their premiums (Donnelley 2014c). This included data on medical histories of patients, identified by date of birth and postcode that could be combined with credit information and lifestyle records. In response the government has announced that it will bring in legislation to protect any records obtained through care.data scheme from being shared for the commercial gain of companies outside the health service (Matthew-King 2014b). Data will only be shared where there is a clear benefit to the NHS and where applicant can demonstrate an 'ethical basis' for its use, although commercial access under these conditions will still be allowed.

5. Discussion

The previous section has presented a detailed picture of three case study projects and organisations involving the large-scale production and use of genomic and/or health data. These present different models of the way in which the relationship between the public and private sector is being constructed in contemporary biomedicine. Whilst a formal comparison is not possible due to fundamental differences in the history, organisation and objectives of these initiatives, it is fruitful to make some reflections on their similarities and differences. This is summarised in Table 1 below.

The first example, UK Biobank, was established in the early 2000s, around the time that new policies associated with the knowledge economy were first being articulated. However, UKBB adopted a well-established organisation form for a large scale research project of this sort, with a clear division between the public and private sector. UKBB is responsible for everything up to data analysis and does generally not engage with industry prior to that point[20]. However, UKBB treats all researchers equally, regardless of their organisational affiliation, and is prepared to send samples out to industrial users (UK Biobank 2011a, p.8). In addition, UKBB has started to spin-out some of its services and expertise into a for-profit enterprise. Its interactions with the private sector may therefore be

[20] With the exception of, for example, renting commercial space to house its assessment centres, employing nurses from agencies, and the aforementioned case of Affymetrix.

characterised as operating in an 'open data' mode in which the emphasis is on enabling private access to the fruits of publicly funded research. To date this has not occurred. However, at the same time the boundary between the public and private sector remains distinct. UKBB cannot in this sense be described as having a market orientation and the creation of a spin-out company, which externalises its commercial activities, merely confirms this.

In contrast, Genomics England and care.data were both established over 10 years later, when policies designed to enable the commercial exploitation of public research (and resources) were well developed. This is reflected in the organisational form and aims of GeL, which integrates the private sector throughout its activities. By comparison with UKBB, GeL is only responsible for overall delivery of the project, data protection and contracting out all other activities, such as sample collection, analysis, and storage to companies, universities, and hospitals. Furthermore, because GeL is seeking to profit ('share the gold') from its activities in a way in which UKBB is not, it is exploring the use of various private forms of collaboration with industry, such as technology licensing and joint ventures. It is therefore governed by a market logic, which is reflected in its legal structure, governance and objectives, but at the same time remains state owned with a public health remit. In this sense, GeL occupies a potentially unstable position within both the bioscience industry and the NHS. This represents a novel form of public-private partnership established by the state, but in which a market orientation is dominant. The uncertainty around its status as neither a research project nor a provider of clinical care reflects this liminal position.

Care.data, the third project described above, is somewhat different from either UKBB or GeL, as organisationally it lies at the heart of the NHS and has been established within a statutory framework as part of a larger co-ordinating body. As with UKBB, and unlike GeL, HSCIC does not generally collaborate with industry directly, but is established to allow analysis of its living data set. Significant private sector use has already been made of the hospital data held by HSCIC, including by a small number of bioinformatics companies. It is therefore reasonable to assume this will continue when care.data comes on line, but within the new legislative framework currently being developed. The main difference between UKBB/ GeL and care.data is that the former are projects that create new data based on the active participation of individuals, whilst the latter uses (pseudonymised) data collected routinely on patients during their interaction with the NHS, but without their active consent. Care.data is therefore best thought of as an enabling open access resource being developed by the state, which will be used by multiple actors for a range of purposes. However, it is clear from the previous analysis of the policy landscape in Section 3 that a major driver behind the creation of care.data is the long-term goal of unlocking the potential of NHS patient information as a valuable asset to be used by the private sector to create new products and services. Strategically, it is a key element in realising the policy vision set out by Poste and Fears in 1999, and reiterated in other policies since, of the NHS as a unique source of untapped economic value.

Social and ethical issues

The relationship between the public and private sector

It is worth at this stage pausing to reflect on why the relationship between the public and private sector should be an issue of concern for policy. The answer to this is rather complex, but in the first instance is a reflection of public concern. A number of surveys, focus groups and other social studies have shown that whilst there is broad support amongst the majority of the population for participation in biomedical research involving the use of personal medical information, this support falls significantly if it involves private companies (Clemence et al, 2013; see Caulfield et al 2014 for a review of the literature in this area).

The focus of many of these studies has been to explore issues of public trust, and their findings suggest that clinicians are seen as highly trustworthy, whereas pharmaceutical companies and the government are generally believed to be less so (Ibid). There are a number of potential explanations for a result which appears to be consistent across time and between different research projects. These include, fears over misuse of participant's bodily samples and data (e.g. secondary use of genetic test results by insurance companies), as well as the general image and reputation of the pharmaceutical industry. However, detailed social science research highlights another important issue relating to the distribution of benefits from research of this sort.

Helen Busby in a study of why people donated blood and gave DNA samples to a genetics related research project (Busby and Martin 2006) found that individuals do not simply give blood or participate in research as a purely unselfish 'altruistic' act. Instead, drawing on anthropological work on gift economies she characterised donation and research participation as belonging to a system of reciprocal exchange. So whilst blood donors did not necessarily expect to receive blood in return during their lifetime, they believed that it might well help their kin or other members of their wider community. Similarly, with biomedical research there was little expectation of direct personal benefit from participation, but a hope that this would help others in the future who belonged to a broader 'imagined community' constituted around the NHS (Ibid). What was striking about the research findings was that this community did not include private firms or people living in other countries, and placed a limit on people's willingness to donate/ participate. In other words, it appears there is a social contract that underpins, and is embedded deeply within, research participation as a system of reciprocal exchange. Making private profit out of this reciprocal relationship appears to breach this contract and the unwritten rules that govern it. Whilst more research needs to be carried out to further elaborate this hypothesis (Kettis-Lindblad et al 2006), it helps shed much deeper insight into why many people are opposed to the use of their health care records and bodily samples to make private profit, and raises important questions about how to conceptualise the proper role of private companies in utilising resources such as NHS medical records.

Participation and consent

Where the basis for participation becomes very important in practical policy terms is in relationship to the issue of consent for the use of an individual's personal medical information and their DNA. One of the key findings from the many studies of the consent process that have been undertaken by social scientists is that the willingness of an individual to give consent to a medical procedure (or to their participation in research) critically depends on the extent to which they trust the individual asking for it. This in turn is very heavily influenced by the institutional context within which consent is sought; for example, whether this is in an NHS clinic or as part of a privately sponsored research project. The content of the specific information given to patients or participants is generally secondary to this in influencing the decision to grant consent.

UKBB has a long track record of seeking consent for participation in its research and found it harder to recruit participants than initially expected. However, this was facilitated by its status as an independent medical research charity primarily studying important diseases affecting many members of the general population. In contrast, during the pilot phase of the 100k Genome Project, GeL is recruiting individuals from patient groups suffering from particular illnesses, such as cancer, rare inherited conditions, or HIV. Whilst these may be easier to recruit than members of the general public, they are also better informed and organised, and a number of critical questions have already been raised by patient groups about the 100k Genome Project during public meetings. It is too early to say if the quasi-private status of GeL will be a barrier to participation, but the analysis presented above suggests that this may be a significant issue in the future unless handled sensitively.

The controversy around care.data further illustrates a number of important points. The guiding policy assumptions behind the decision to pursue an opt-out model appears to be that this is already well established custom and practice for use of anonymised data within public health research. The adoption of an opt-out was largely a concession in the face of criticism, but did not fundamentally change the ethos of the project. What has become problematic, however, is the disputed status of the data as potentially identifiable and therefore requiring a different form of consent and governance. This has been exacerbated by official concerns about data security, the acknowledged major shortcomings of the public information campaign, and the sale of hospital data to the insurance industry. As a result there has been a significant loss of public trust in the management and benefits of the project.

In this context, there is a real danger that policy makers will continue to assume that the opposition to care.data can be overcome by improving public relations and drawing on the idea that the NHS embodies a form of social solidarity that still command high levels of support. However, as this analysis has shown, this would fundamentally misunderstand the nature of the problem. There are profound reasons why there is so little support for the use of patient data of this sort without consent, where private companies will benefit, and when people feel they cannot trust the state.

Anonymity and data sharing

A well-worn set of concerns surrounding the creation of large databases of the sort described here focus on the question of anonymity. In UKBB participant information is anonymised, but can still be re-identified to enable new data to be linked to existing records. In the case of GeL the main issue of concern is that whole genome sequences can be used to identify individuals and therefore must be stored behind robust firewalls. Finally, in the case of care.data, potentially sensitive medical information held on individuals and which could be used to identify them, will be kept in a pseudonymised form. In each project robust data security and governance procedures are being established to protect privacy and ensure confidentiality.

Leaving aside the real concerns about the practicalities of doing this and the sometimes poor track record of the security in large public databases, private access to these resources remains a major issue. In particular, serious thought needs to be given to the kinds of oversight, regulation and incentives that will prevent companies from seeking to re-identify individuals. This may involve the use of specific contractual condition and penalties, novel institutional arrangements such as safe harbours, but also the development of a much deeper culture of care and respect for the data provided by patients and participants.

Expectations of future benefits

One of the most important assumptions about each of these three projects is that they will bring direct benefits to patients, improve public health and stimulate economic growth. As described in Section 3, the idea that developing large infrastructures of this sort could be a source of economic value was something that only became part of the mainstream policy discourse in the last 15 years.

In terms of health, for the advocates of genomics many of these benefits will be realised in the not too distant future. For example, the report by the influential Human Genomics Strategy Group (2012), which made the case for an initiative along the lines of the 100k Genome Project claimed that:

> "We are currently on the cusp of a revolution in healthcare: genomic medicine – patient diagnosis and treatment based on information about a person's entire DNA sequence, or 'genome' – becoming part of mainstream healthcare practice." (p.14)
>
> "And, with the pace of technological change at an unprecedented level, many uses of genomics are set to enter mainstream clinical practice in the next three to five years." (p.17)

However, there are good reasons to be cautious about uncritically accepting some of these claimed benefits and the timescale they will be delivered in. The history of genomics has been marked by high expectations that have had to be reduced significantly over time (Hopkins et al 2007). For example, few gene variants have been found to directly cause common complex diseases, as was originally anticipated, and few have formed the basis for widely used diagnostics (Wolf et al 2009). Furthermore, it has proved extremely hard to translate the rapid advances in basic science into techniques and technologies that are widely used in routine clinical practice. This is clearly shown in the very small number of drugs on the market that are directly derived from genomic targets (Isserlin et al 2011). The same is true for gene-based diagnostics relating to common complex disorders, something which was widely expected during the 1990s and early 2000s.

In terms of the general economic benefits of creating large data resources, a recent report by Capgemini Consulting (Tinholt 2013: 7) claims that Big/Open Governmental Data (OGD) relating to public sector organisations can firstly bring economic benefits by enabling private companies to develop new products and services as described above, and thus create jobs and generate tax revenue. In addition, the involvement of the private sector may allow costs to be cut and service improvements to be made within the public sector. Two examples are frequently drawn upon to demonstrate the benefit of such private/public relations in healthcare. Firstly, the discovery that more patients undergoing elective surgery within the NHS die at the weekend compared to other days of the week was identified by a private company, Dr Foster Intelligence, following analysis of data made available by HSCIC (Aylin et al. 2013). A second finding, that certain regions in the UK are more likely to prescribe expensive statins, was discovered by private company Mastodon-C, incubated by the Open Data Institute and Open Health Care UK (Prescribing Analytics 2012). Genomics England anticipate similar benefits to emerge following analysis of their 100k Genome Project. It claims this will enable healthcare professionals to correctly identify disease aetiology, allowing them to go "straight to the answer [the correct diagnosis] rather than spending 20 years wandering around the hospital wards." (Genomics England 2013f). How quickly this might be achieved remains unclear.

In considering the more specific commercial benefits of genomics, it is worth noting that the biotechnology industry is marked by very long products lead times, a high rate of failure and few companies that are financially sustainable through product sales. This is not to say that the industry will not be successful in the long-term, but that as described in Section 3 there has been significant consolidation and retrenchment of the genomics sector in recent years, making its potential short-term contribution to the UK economy from involvement in these projects limited.

These points are not made in order to reject the very real promise that genomics holds for the long-term development of technologies that will significantly improve patient care and contributing to the foundation for a successful bioscience industry. Rather, they are given as a note of caution about the expectations that a revolution in genomic medicine is just around the corner. Instead, a different model of innovation based on detailed empirical evidence is required to guide policy thinking, in which public research is not always driven by the perceived short-term demands of the market for research that has direct economic impact. There is a substantial body of scholarship within innovation studies that highlights the incremental nature of technological change in medicine (see Hopkins et al 2007) and the importance of supporting blue skies basic research for long-term economic growth (Mazzucato 2014).

What is in the public interest?

Finally, it is worth asking a question about what is the public interest? At the heart of the policy consensus described earlier is the assumption that the private sector and market mechanisms are the most efficient way of both improving individual and public health, whilst at the same time leading to economic growth. The example of Genomics England highlights the way in which this is now being enshrined within public policy, so that the state increasingly uses market-based approaches to deliver public benefits through enabling the private sector to develop new goods and services.

However, as discussed above, it is far from certain that many of the claimed benefits of the three projects for improving both the health and wealth will be delivered in the near future. If this is the case, then a different rationale has to be used to justify the creating of large and costly infrastructures, which is about strategic investment for national benefit. This cannot be simply guided by short-term private economic interests, but also has to have a strong mandate for long-term improvement in public health. The key issue therefore is the extent to which a market-oriented approach to the design and delivery of policy can achieve this.

6. Has the relationship between the public and private sector changed?

In conclusion, we can return to the question outlined at the start of the report and attempt to assess the extent to which the relationship between the public and private sector has changed in the field of bioscience both over the last 30 years, as well is in more recent times.

This report has argued that a major shift in the dominant economic paradigm occurred in the 1990s and 2000 which saw the rise of a new set of ideas about the role of knowledge and technology in stimulating economic growth and competition. One major consequence of this shift to a knowledge economy were moves to restructure the relationship between the public and private sector, with academia and public research increasingly seen as creating new knowledge that could be commercially exploited by industry. Within this framework the priority has been to ensure the efficient transfer and exchange of knowledge between the public and private sector in order to enable knowledge-based innovation. This has had a major impact in the funding, governance and organisation of public biomedical research programmes.

More recently, a related discourse has started to gain increasing prominence that reimagines the NHS as a source of economic value. This has taken a number of forms, most notably a vision of the commercial benefits that might be gained from the secondary analysis of linked genomic and patient information via large IT infrastructures. The extraction of private value from public data of this sort is increasingly justified under the new banner of Big Data, but in this case precedes it by a number of years.

Despite the changes outlined above, it must be stressed that the evidence presented in this report suggests that there has not been a fundamental change in the scale of direct engagement by industry with the public sector in this area. This is particularly true of large pharmaceutical companies and dedicated biotech firms, who remain largely absent from the projects we have examined. This is partly a result of the restructuring and consolidation of the genomics industry, but also underlies the extent to which these projects are mainly strategic and enabling, rather than near market in their application.

However, what is apparent from this study is the way in which the nature of the public sector itself is changing with increasing use of business-type organisations and methods. The commitment to the alignment and integration of the public and private sector in the name of knowledge exchange and value extraction seen in the case of GeL, and to a lesser extend care.data, can be thought of as a novel form of public-private partnership. It involves major public investment in skills and infrastructure to enable private access to personal medical and genomic data on a huge scale, but in a framework in which the direction of public research is increasingly subordinated to the perceived needs of industry. This raises many questions about the goals of public investment, the steering and direction of research, the distribution of benefits, and the governance of research.

Table 1: Summary of key features of case studies

	UK Biobank	**Care.data**	**Genomics England**
Organisation	Independent organisation	Project undertaken by statutory body	Business established to deliver 100k Genome Project
Legal status	Not-for-profit charity	Part of HSCIC - Executive Non Departmental Public Body	For profit private limited company owned by the Department of Health
Mandate	Research	Provision of information to improve care & enable research	To be clarified
Funding	Government & charitable	Government	Government
Data held	Genotype, biomedical test results, and personal medical information	Personal medical information	Whole genome sequence
Consent	Yes – broad consent covering future use of data	No - but possibility of opting out of database	Yes – broad consent covering future use of data
Anonymity	Anonymised, but participants can be identified as it is a prospective study	Pseudanonymised – possible to identify individuals	Anonymised, but concerns that not possible in practice with WGS data
Data sharing	Data will be linked with various NHS data sets	Integration of hospital & GP records. Linkage to other NHS data sets	Linkage to care.data
Model of public-private partnership	A research resource. No formal collaboration with private sector. Data made available for research	'Opening up' of data held by statutory body to enable innovation	A private company funded by state to undertake 100K Genome Project. Maximise revenue & health benefits
Role of private sector in delivering project	Firms involved in genotyping. Creation of spin-out BioCentres to provide for-profit services	Firms involved in some aspects of creating database (GPES)	Contracting out of all main technical activities
Private sector access to data	Yes - for nominal admin fee	Yes – for nominal admin fee	Yes – on commercial terms (licensing, joint venture, fee)
Rationale for private access	To undertake biomedical research	To improve patient care through analysis of dataset	To foster innovation & generate revenue for GeL
Use of data by private sector (by April 2014)	None (since March 2012)	Yes – mainly bioinformatics firms	Resource under construction

6. References

Anon, 2013. *Healthcare Evaluation Data (HED)*. , p.2. Available at: http://www.uhb.nhs.uk/Downloads/pdf/HedIntroLeaflet.pdf.

Aylin, P. et al., 2013. Day of week of procedure and 30 day mortality for elective surgery: Retrospective analysis of hospital. *BMJ*, 346(May), pp.2424–2432.

Bains, W (2004) Failure rates in drug discovery and development: will we ever get any better? *Drug Discovery World*. Fall 2004 p9-18

Bhatia, N., (2014) *care.data*. Available at: http://www.care-data.info/ [Accessed April 8, 2014].

Bioscie3nce Innovation and Growth Team (BIGT) (2003) *Bioscience 2015: Improving National Health, Increasing National Wealth*. A Report to Government by the Bioscience Innovation and Growth Team. London: Departmet for Trade and Industry. Avaialble from http://www.bioindustry.org/document-library/bioscience-2015/ Accessed July 2014.

Boliver, V., 2010. Expansion, differentiation, and the persistence of social class inequalities in British higher education. *Higher Education*, 61(3), pp.229–242. Available at: http://link.springer.com/10.1007/s10734-010-9374-y [Accessed January 10, 2014].

Busby, H. and Martin, P.A. (2006) Biobanks, national identity and imagined genetic communities: the case of UK biobank. *Science as Culture.* 15(3):237-251.

Business Wire, 2013. *University of Cambridge, Genomics England Ltd., and Illumina collaborate to sequence 100,000 whole human genomes for National Health Service* patients. Available at: http://www.businesswire.com/news/home/20131021005458/en/University-Cambridge-Genomics-England-Ltd.-Illumina-Collaborate#.U0PE4PldV8F [Accessed April 8, 2014].

Caulfield, T. (2014) A review of the key issues associated with the commercialization of biobanks. *Journal of Law and the Biosciences*. 1 (1): 94-110.

Chowdry, H. et al., (2008) *Widening participation in higher education: Analysis using linked administrative data*, London.

Clemence, M. et al., (2013) *Wellcome Trust Monitor Wave 2: Tracking Public Views on Science, Biomedical Research and Science Education*. Social Research Institute: London

Cowan, R. and van de Paal, G. (2000) *Innovation Policy in a Knowledge-Based Economy.* A MERIT Study commissioned by the European Commission Enterprise DirectorateGeneral. Publication no. EUR 17023 of the Commission of the European Communities, Luxembourg. Avaialble from ftp://ftp.cordis.europa.eu/pub/innovation-policy/studies/studies_knowledge_based_economy.pdf. Accessed 14.4.2014

Dawson, C.R., (2000). *Public perceptions of the collection of human biological samples.* , p.134. Available at: https://www.ukbiobank.ac.uk/wp-content/uploads/2011/07/Public-Perceptions-Collection-Human-Biological-Samples.pdf?phpMyAdmin=trmKQlYdjjnQIgJ%2CfAzikMhEnx6.

Department of Health (2012) *Modernising health and care public bodies - The Health and Social Care Act.* , (April), p.2. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138265/B6.-Factsheet-Streamlined-arms-length-bodies-250412.pdf.

Department of Health (2003) *Our Inheritance, Our Future: Realising the potential of genetics in the NHS*, www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4019239.pdf

Donnelley, L. (2014a) NHS admits new medical records database could pose privacy risk. *The Telegraph*. 16th February 2014. Available from: http://www.telegraph.co.uk/health/nhs/10642740/NHS-admits-new-medical-records-database-could-pose-privacy-risk.html. Accessed 25.4.14.

Donneley, L. (2014b) NHS blunders put millions of records at risk. *The Telegraph.* 8th April 2014. Available from http://www.telegraph.co.uk/health/healthnews/10753001/NHS-blunders-put-millions-of-records-at-risk.html. Accessed 25.4.14.

Donnelley, L. (2014c) Hospital records of all NHS patients sold to insurers. *The Telegraph*. 27th February 2014. Available from: http://www.telegraph.co.uk/health/healthnews/10656893/Hospital-records-of-all-NHS-patients-sold-to-insurers.html. Accessed 25.4.14.

Etzkowitz, H. & Leydesdorff, L., (2000) The dynamics of innovation: From national systems and "'Mode 2'" to a triple helix of university–industry–government relations. *Research Policy*, 29, pp.109–123.

Epstein, S., (1996) *Impure Science: AIDS, Activism and the Politics of Knowledge*, Berkeley: University of California Press.

Epstein, S., (1995) The construction of lay expertise: AIDS and the forging of credibility in the reform of clinical trials. *Science, Technology & Human Values*, 20(4), pp.408–437.

Fears, R. and Poste, G. (1999) Building population genetics resourcesusing the UK NHS. *Science* 284(5412) pp267-268.

Felt, U. et al (2007) *Taking European Knowledge Society Seriously.* Directorate-General for Research, Science, Economy and Society. European Comission. Available at: https://www.bmbf.de/pub/EuropeanKnowledge(6).pdf [Accessed June 2014]

Frank, J. et al., (2006) Large life-course cohorts for characterizing genetic and environmental contributions: the need for more thoughtful designs. *Epidemiology*, 17(6), pp.595–598.

Genomics England (2013a) *About Genomics England*. Available at: http://www.genomicsengland.co.uk/100k-genome-project/ [Accessed March 31, 2014].

Genomics England (2014a) *Ethics and genomics survey.* Available at: http://www.genomicsengland.co.uk/ethics-and-genomics-survey.html [Accessed April 8, 2014].

Genomics England (2013b) *Genomics England annotator supply event. Part 1.*, Available at: http://www.genomicsengland.co.uk/annotation-supplier-event/.

Genomics England (2013c) *Genomics England annotator supply event. Part 2.*, Available at: http://www.genomicsengland.co.uk/annotation-supplier-event/.

Genomics England, (2013d) Genomics England annotator supply presentation. , p.25. Available at: http://www.genomicsengland.co.uk/wp-content/uploads/2013/11/Annotation-Supplier-Event-November-5-2013.pdf.

Genomics England (2013e). *Genomics England town hall engagement event. Afternoon: For clinicians, scientists, healthcare professionals, funders and allied organisations. Part 1.*, Available at: http://www.genomicsengland.co.uk/town-hall-engagement-event/.

Genomics England (2013f) *Genomics England town hall engagement event. Afternoon: For clinicians, scientists, healthcare professionals, funders and allied organisations. Part 2.*, Available at: http://www.genomicsengland.co.uk/town-hall-engagement-event/.

Genomics England (2013g) *Genomics England town hall engagement event. Morning: for public, patients and patient support charities. Part 1.*, Available at: http://www.genomicsengland.co.uk/town-hall-engagement-event/.

Genomics England (2013h) *Genomics England town hall engagement event. Morning: for public, patients and patient support charities. Part 2.*, Available at: http://www.genomicsengland.co.uk/town-hall-engagement-event/.

Genomics England, 2014b. How we work. Available at: http://www.genomicsengland.co.uk/how-we-work/ [Accessed April 14, 2014].

Genomics England (2013i) S*oftware innovators being invited to bid for share of £10 million fund.* Available at: http://www.genomicsengland.co.uk/software-innovators-being-invited-to-bid-for-share-of-10-million-fund/ [Accessed April 8, 2014].

Gibbons, M. et al., (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*, London: Sage Publications.

Health and Social Care Information Centre (2014a) *A strategory for the Health and Social Care Information Centre 2013-2015*. , p.48. Available at: http://www.hscic.gov.uk/media/13557/A-strategy-for-the-Health-and-Social-Care-Information-Centre-2013-2015/pdf/hscic-strategy-2014.pdf.

Health and Social Care Information Centre (2014b) *Data Access Advisory Group (DAAG)*. Available at: http://www.hscic.gov.uk/daag [Accessed April 8, 2014].

Health and Social Care Information Centre (2013a). *Data linkage and extract service. Service charges 2013/14*. , p.4. Available at: http://www.hscic.gov.uk/media/12443/data-linkage-service-charges-2013-2014-updated/pdf/dles_service_charges__2013_14_V10_050913.pdf.

Health and Social Care Information Centre (2012) *Exploiting the potential of health and care data*, Available at: http://www.hscic.gov.uk/ourrole.

Health and Social Care Information Centre (2014c). Hospital episode statistics. Available at: http://www.hscic.gov.uk/hes [Accessed April 8, 2014].

Health and Social Care Information Centre (2014d). HSCIC agrees improvements to how information is shared. Available at: http://www.hscic.gov.uk/article/3952/HSCIC-agrees-improvements-to-how-information-is-shared [Accessed April 8, 2014].

Health and Social Care Information Centre (2013b). Obtaining consent for patient identifiable and/or sensitive data. , p.2. Available at: http://www.hscic.gov.uk/media/12053/Obtaining-consent-for-patient-identifiable-and-or-sensitive-data-V2-050613/pdf/obtaining_consent_for_patient_identifiable_and_or_sensitive_data_V2_050613.pdf .

Health and Social Care Information Centre (2014e). Supporting transparency and open data. Available at: http://www.hscic.gov.uk/transparency [Accessed April 8, 2014].

Health and Social Care Information Centre & NHS England (2014). Frequently asked questions for patients. , p.8. Available at: http://www.england.nhs.uk/wp-content/uploads/2014/02/cd-pati-faq240214.pdf.

Health and Social Care Information Centre & NHS England (2013). How information about you helps us to provide better care. Frequently asked questions for patients. , p.7. Available at: http://www.nhs.uk/NHSEngland/thenhs/records/Documents/Patient FAQs.pdf.

Healthcare Industry Task Force (HITF) (2004) B*etter Healthcare through partnership: a programme for action*. London: Department of Health. Available from http://webarchive.nationalarchives.gov.uk/20040809170550/http://advisorybodies.doh.gov.uk/hitf/newsdesk.htm. Accessed July 2014.

Hessels, L.K. & van Lente, H., 2008. Re-thinking new knowledge production: A literature review and a research agenda. *Research Policy*, 37(4), pp.740–760

Hopkins, MH., Martin, PA., Nightingale, P., Kraft, A. and Mahdi, S. (2007) The Myth of the Biotech Revolution: An assessment of technological, clinical and organisational change. *Research Policy **36**(4): 566-589*

Human Genomics Strategy Group (2012) *Building on our inheritance: genomic technology in Healthcare.* London: Department of Health. Available form: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213705/dh_132382.pdf Accessed July14

Isserlin, R. et al (2011) The human genome and drug discovery after a decade. Roads (still) not taken. *Nature* 470, 163-165

Keogh, B. & Kelsey, T., (2013) NHS England Genomics Strategy Board. , p.8. Available at: http://ukgtn.nhs.uk/fileadmin/uploads/ukgtn/Documents/Resources/Library/External_Documents/NHS England Genomics Strategy Board.pdf.

Kettis-Lindblad, A. et al., (2006) Genetic Research and Donation of Tissue Samples to Biobanks. What Do Potential Sample Donors in the Swedish General Public Think? *Eur. J. Public Health* 433

Lewis, G., (2014) Better information means better care - Dr Geraint Lewis. Available at: http://www.england.nhs.uk/2014/01/15/geraint-lewis/ [Accessed April 8, 2014].

Lewis, G., (2013) Introducing the care.data programme. , (July), p.12. Available at: http://www.hscic.gov.uk/media/12350/stakeholder-forum-10-07-13-5introducing-the-care-data-programme/pdf/Stakeholder_Forum_10-07-13-5_-_Introducing_the_care.data_Programme.pdf.

Malerba, F. and Orsenigo, L. (2002) Innovation and market structure in the dynamics of the pharmaceutical industry and biotechnology: towards a history-friendly model. *Industrial and Corporate Change.* 11 (4) pp667-703.

Martin, P. and Kaye. J. (1999) *The Use of Biological Sample Collections and Personal Medical Information in Human Genetics Research*. 64 pages. Wellcome Trust, London.

Matthews-King, A. (2014a)Three in four GPs believe care.data should be 'opt in'. *Pulse Magazine*. 26 February 2014. Available from: http://www.pulsetoday.co.uk/your-practice/practice-topics/it/three-in-four-gps-believe-caredata-should-be-opt-in/20005954.article#.U1ozxvldXSg. Accessed 25.4.14.

Matthews-King, A. (2014b) Legislation will prevent patient data being sold outside the health service, says Hunt. *Pulse Magazine*. 3 March 2014. Available from: http://www.pulsetoday.co.uk/your-practice/practice-topics/it/legislation-will-prevent-patient-data-being-sold-outside-the-health-service-says-hunt/20006014.article#.U1oxRPldXSg. Accessed 25.4.14

Mazzucato, M. (2014) *The entrepreneurial state: debumkimg public vs private sector myths*. London: Anthem Press

Mowery, D.C. , Nelson, R.R., Sampat, B.N., Ziedonis, A.A. (2001) The growth of patenting and licensing by U.S. universities: an assessment of the effects of the Bayh–Dole act of 1980. *Research Policy*. 30 (1), pp99–119

Narayanan, A. & Shmatikov, V., (2008) Robust de-anonymization of large datasets. *Security and Privacy*, pp.111–125.

NHS England, (2014). Better information means better care. Available at: http://www.nhs.uk/NHSEngland/thenhs/records/healthrecords/Pages/care-data.aspx [Accessed April 8, 2014].

NHS (2011) *Innovation, Health and Wealth: Accelerating adoption and diffusion in the NHS.* London NHS England. Available from: http://www.institute.nhs.uk/images/documents/Innovation/Innovation%20Health%20and%20Wealth%20-%20accelerating%20adoption%20and%20diffusion%20in%20the%20NHS.pdf Accessed July 2014

Nowotny, H., Scott, P. & Gibbons, M., (2003) "Mode 2" revisited: The new production of knowledge. *Minerva*, 41(3), pp.179–194.

Nowotny, H., Scott, P. & Gibbons, M., (2001). *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*, Cambridge, UK: Polity Press.

Owen-Smith, J. and Powel W.W. (2001), Careers and contradictions: faculty responses to the transformation of knowledge and its uses in the life sciences, in Steven Vallas (ed.) *The Transformation of Work* (Research in the Sociology of Work, Volume 10) Emerald Group Publishing Limited, pp.109 - 140

Palin, M., (2013). Mark Palin... welcomes the Personal Genome Project and highlights key differences to GeL's 100K Genome Project. Available at: http://www.genomicsengland.co.uk/mark-palin-welcomes-the-personal-genome-project-and-highlights-its-key-differences-to-gels-100k-genome-project/ [Accessed April 8, 2014].

Parker, M., (2013a) No Title. , p.7. Available at: http://www.genomicsengland.co.uk/wp-content/uploads/2013/06/GenomicsEngland_EthicsWorkingGroup.pdf.

Parker, M., (2013b). Prof Michael Parker... says no ethical issues are off the table as GeL considers its approach to patient consent. Available at: http://www.genomicsengland.co.uk/prof-mike-parker-says-no-ethical-issues-are-off-the-table-as-gel-considers-its-approach-to-patient-consent/ [Accessed April 8, 2014].

Pisano, G. (2006) *Science Business*. Cambridge MA: Harvard Business School Press

Prescribing Analytics, 2012. Methodology. Available at: http://www.prescribinganalytics.com/methodology [Accessed December 12, 2014].

Richards T (2000) Court sanctions use of anonymised patient data. *British Medical Journal*, 320, 77)

Royal College of General Practitioners et al., (2014) Frequently asked questions - care.data guide for GP practices. , p.10. Available at: http://www.england.nhs.uk/wp-content/uploads/2014/03/cd-gp-faq-03-14.pdf.

Shinn, T., (2002) The triple helix and new production of knowledge: Prepackaged thinking on science and technology. *Social Studies of Science*, 32(4), pp.599–614.

Tinholt, D., 2013. *The Open Data Economy Unlocking Economic Value by Opening Government and Public Data.*

Triggle, N. (2014) *Care.data: How did it go so wrong?* BBC Health News Available from http://www.bbc.co.uk/news/health-26259101 Accessed April 2014

UK Biobank (2014a) About UK Biobank. Available at: http://www.ukbiobank.ac.uk/about-biobank-uk/ [Accessed April 8, 2014].

UK Biobank, 2011a. Access procedures: Application and review procedures for access to the UK Biobank Resource. , 0(November), p.36. Available at: http://www.ukbiobank.ac.uk/wp-content/uploads/2012/09/Access-Procedures-2011.pdf.

UK Biobank, 2011b. Public consultation on draft access procedures: Summary of responses and modifications. , (September 2011), p.11. Available at: https://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Consultation-Report-210911.pdf?phpMyAdmin=trmKQlYdjjnQIgJ%2CfAzikMhEnx6.

UK Biobank, 2014h. UK Biobank board. Available at: http://www.ukbiobank.ac.uk/uk-biobank-board/ [Accessed December 12, 2014].

UK Biobank, 2007a. UK Biobank ethics and governance framework version 3.0. , 0(October), p.20. Available at: https://www.ukbiobank.ac.uk/wp-content/uploads/2011/05/EGF20082.pdf?phpMyAdmin=trmKQlYdjjnQIgJ%2CfAzikMhEnx6.

UK Biobank, 2014b. UK Biobank opens for research. Available at: http://www.ukbiobank.ac.uk/2012/06/uk-biobank-opens-for-research/ [Accessed April 8, 2014].

UK Biobank, 2013. UK Biobank re-contact procedures. , p.13. Available at: http://www.ukbiobank.ac.uk/wp-content/uploads/2014/02/Re-contact-Procedures-22-1-14.pdf.

UK Biobank, 2014c. UK Biobank showcase user guide: Getting started. , (February), p.8. Available at: http://biobank.ctsu.ox.ac.uk/~bbdatan/ShowcaseUserGuide.pdf.

UK Biobank, 2014d. UK Biobank welcomes imaging funding. Available at: http://www.ukbiobank.ac.uk/2013/09/uk-biobank-welcomes-imaging-funding/ [Accessed April 8, 2014].

UK Biobank, 2007b. UK Biobank: Protocol for a large-scale prospective epidemiological resource. , 06(March), p.112. Available at: http://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf?phpMyAdmin=trmKQlYdjjnQIgJ%2CfAzikMhEnx6.

UK Biobank, 2014e. What diseases will UK Biobank help scientists study? Available at: https://www.ukbiobank.ac.uk/faqs/what-diseases-will-uk-biobank-help-scientists-study/ [Accessed April 8, 2014].

UK Biobank, 2014f. What happens if a researcher makes a profit from using the Resource? Available at: http://www.ukbiobank.ac.uk/faqs/what-happens-if-a-researcher-makes-a-profit-from-using-the-resource/ [Accessed April 8, 2014].

UK Biobank, 2014g. Will tobacco industry researchers be able to use the Resource. Available at: http://www.ukbiobank.ac.uk/faqs/will-tobacco-industry-researchers-be-able-to-use-the-resource/ [Accessed April 8, 2014].

UK Biobank Interim Advisory Group on Ethics and Governance (2003) UK Biobank ethics and governance framework background document. , p.24. Available at: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@msh_grants/documents/web_document/WTD003287.doc.

UK Biocentre (2013) Case study: UK Biocentre sample processing and storage for the University of Cambridge on behalf of the National Health Service Blood and Transplant INTERVAL Initiative. , p.2.

UK Biocentre, (2014) UK Biocentre. Available at: https://www.ukbiobank.ac.uk/uk-biocentre-2/ [Accessed April 8, 2014].

Wolf, H., et al (2009) Challenges of translating genetic tests into clinical and public health practice. *Nature Reviews Genetics*. 10, pp489-495

Wright, C. et al., (2011) *Next steps in the sequence: The implications of whole genome sequencing for health in the UK*, Cambridge, UK.