Applied & Translational Genomics 3 (2014) 116–119

Contents lists available at ScienceDirect

Applied & Translational Genomics

journal homepage: www.elsevier.com/locate/atg

Review

To share or not to share is the question

Donald R.C. Chalmers *, Dianne Nicol, Margaret F. Otlowski

Centre for Law and Genetics, University of Tasmania, Private Bag 89, Hobart, TAS 7001, Australia

ARTICLE INFO

Keywords:
Data sharing
Personalised medicine
Privacy

ABSTRACT

Data sharing is increasingly becoming an essential component of clinical practice and biomedical research. The debate has shifted from whether or not to exchange data to how best to achieve optimal sharing. This raises new ethical and legal challenges, particularly with regard to consent and privacy. This article discusses recent developments in the formulation of best practice guidelines for data sharing. Particular attention is focused on the Global Alliance for Genomics and Health (GA4GH) draft Framework of Conduct for Data Sharing.

Contents

1. Introduction

As the power of integrating multiple sources of data to progress understanding of human health is becoming increasingly understood, there is a general recognition that the next phase of personalised medicine will see acceleration in data sharing to link genome scans to clinical data.[1] In the clinical context, the move to electronic health records and electronically stored data provides opportunities to use and share data to better understand disease and illness, inform treatment choices and patient care and improve health outcomes.[2] In the research context, organisations such as the International Cancer Genome Consortium (ICGC)[3] and, more recently, the Global Alliance for Genomics and Health (GA4GH)[4] have embraced policies and plans to proselytise and promote the exchange of clinical data not only amongst Consortium and Alliance members but also more widely in research and clinical care. The overarching aim is to drive the research into the translation phase where the clinical data will be matched with genomic data to inform the development of treatments and medications.

'Data sharing' can take many different forms; e.g. patients agreeing to share their genomic and/or clinical data with researchers; researchers sharing their preliminary data with other researchers; biobanks and other holders of specimens and data sharing their resources with researchers in other countries. This paper encompasses all such forms of data sharing, but is particularly focused on larger scale data sharing involving multiple players, across jurisdictions.

The importance of data sharing has become something of a 'mantra'[5] amongst medical and health researchers. This mantra has been fashioned by government initiatives to promote the new knowledge economy.[6] As an example in the clinical context, the Strategy for UK Life Sciences states:

* Corresponding author. Tel.: +61 362267567.

[1] Knoppers BM et al (2011). 'Towards a Data Sharing Code of Conduct for International Genomic Research', *Genome Medicine* 3:46–49; Kaye J et al (2009) 'Data Sharing in Genomics — Re-shaping Scientific Practice', *Nature Genetics* 10(5):331–335.

[2] Schneeweiss, S (2014). 'Learning from Big Health Care', *New England Journal of Medicine* 370(23):2161–2163; Olver I, (2014) 'Linking Data to Improve Health Outcomes' 200 *Medical Journal of Australia* 368.

[3] http://icgc.org.

[4] http://genomicsandhealth.org/.

[5] Jasny B (2013). 'Realities of Data Sharing Using the Genome Wars as a Case Study — an Historical Perspective and Commentary', *EPJ Data Science* 2:1 doi:10.1140/epjds13.

[6] Academy of Medical Sciences UK (2010), *Review of the Regulation and Governance of Medical Research*, http://www.acmedsci.ac.uk/.

http://dx.doi.org/10.1016/j.atg.2014.09.011

'The UK can do much more to harness the opportunity that exists in the NHS. There is huge potential to better support the adoption and diffusion of innovation, to access patient data to inform the developmental phase, and to involve patients in trials and early access schemes for the treatment of chronic diseases, such as cancer.'[7]

The sheer volumes of data are creating mountainous storage and download exchange challenges.[8] The debate has shifted from whether or not to share data for research, together with the accompanying technical challenges, to how data exchange should be done in order to add value to the research endeavour whilst protecting research participants.[9] However, like genomic science itself, data sharing in this arena has not proven to be easy and a long road lies ahead. There is recognition of the need for a risk/benefit analysis; whilst data sharing is seen as essential to promote the goals of the genome era, care must be taken to minimise the risk of harm from such data sharing. Of its nature, genetic data has some particular characteristics: genetic information is ubiquitous, permanent and unalterable. Even when de-identified, genetic data is always inherently identifiable, and this applies also to person's whole genome sequence, so special protections are required if such data is to be linked to other sensitive information.[10] As mechanisms for data protection become increasingly sophisticated, risk arises from new strategies to outflank protections.[11] Once data is released into the public domain, neither participants nor researchers can control its use, or the possibility of that data being linked to other data sets.[12] The pitfalls for data sharing are many, with privacy, industry–academia divides, distinction between first and third world technological capacities, and diverse researcher, clinical and institutional practices amongst the regulatory hurdles across national borders. [13] Other challenges include workforce and infrastructure limitations but one of the greatest challenges is overcoming policy issues.[14] Kaye has identified four particular areas for attention:

'The difficulties of acknowledging individual contributions to the generation of data; the way that these policies change the responsibilities towards participants; the implications that this has for maintaining public trust; and the new mechanisms that have been developed for oversight of access to data.'[15]

A key issue in this context is the level of informed consent for data sharing. Potentially there are a range of models – at one extreme – no consent or notice, or notice only, or 'opt-out' rights or 'opt in' rights or other forms of express consent. Because of the scale of genomic data and very nature of biobanks as platforms for research undertaken over a period of time, there has been considerable support for a 'broad consent' model whereby participants give agreement to the use of their samples and information, in a de-identified form, for future as yet unspecified research, subject to normal ethics committee review and this approach is endorsed in a number of jurisdictions.[16] There are some commentators, however, who contest that this can ever be an effective consent.[17]

Kaye has promoted technologically aided 'dynamic consent' as part of a more sophisticated genomic data management system: i.e. 'a personalised, digital interface that connects researchers and participants,' facilitating 'two-way communication to stimulate a more engaged and informed.. participant population where individuals can tailor and manage their own consent and preference.'[18] There is continuing debate about optimal consent models for biobanks and large data sharing platforms[19]; for the purposes of this paper, as a minimum, broad consent should be obtained from participants to future genomic research and for data sharing, as a precondition for data sharing.

The development of data sharing policies and practices will require the development of standards.[20] This article examines how far an international code or framework of ethics may contribute to changing attitudes and practices towards more responsible and secure sharing of research and clinical data.

2. Background to data sharing

There has been major expansion of globalisation of research in the 'Genome Era'. This has prompted a range of international organisations to enter the arena of international ethics standard setting. Examples are UNESCO (and their trilogy of Declarations on the Human Genome and Human Rights, 1997; Human Genetic Data 2005; and Bioethics and Human Rights 2005) and the OECD (particularly their Report on *Creation and Governance of Human Genetic Research Databases*, in 2007 and their *Guidelines on Human Biobanks and Genetic Research Databases* 2009). Similarly, there has been a great deal of progress by national organisations in the development of governance frameworks for biorepositories, which are seen as essential resources for global genomic research. As examples, the National Cancer Institute of the National Institutes of Health provided guidance on biobanks in 2006,[21] as did the international Human Genome Organisation in their Human Genomic Databases Report in 2002. In Australia, the National Health and Medical Research Council commissioned a Biobanks Information Paper in 2010.

From the 1996 Human Genome Project Bermuda Declaration onwards, researchers themselves have also embraced the data sharing movement.[22] There is a realisation by researchers of the power of shared data.[23] This can, however, represent a tension with university policies focusing on protection of intellectual property rights, engagement with industry and formalisation of exchanges of materials.[24] Despite this, it is widely understood that genomic research is a

[7] Academy of Medical Sciences UK (2010), *Review of the Regulation and Governance of Medical Research*, http://www.acmedsci.ac.uk/.

[8] Google, Amazon and Microsoft are active in this new cloud commercial environment.

[9] Ohno-Machado, L (2012) 'To Share or not to Share: That is Not the Question' Science Translation Medicine 4:1–4.

[10] Ohno-Machado, L (2012) 'To Share or not to Share: That is Not the Question' Science Translation Medicine 4:1–4.

[11] Erlich Y and Narayanan A, (2014) 'Routes for Breaching and Protecting Genetic Privacy' *Nature Review Genetics* 15:409.

[12] Kaye J et al (2009). 'Data Sharing in Genomics — Re-shaping Scientific Practice', *Nature Genetics* 10(5):331–335.

[13] Jasny B (2013). 'Realities of Data Sharing Using the Genome Wars as a Case Study — an Historical Perspective and Commentary', *EPJ Data Science* 2:1 doi:10.1140/epjds13.

[14] Ohno-Machado, L (2012) 'To Share or not to Share: That is Not the Question' Science Translation Medicine 4:1–4.

[15] Kaye J et al (2009). 'Data Sharing in Genomics — Re-shaping Scientific Practice', *Nature Genetics* 10(5):331–335.

[16] Otlowski, M (2012). 'Tackling Legal Challenges Posed by Population Biobanks: Reconceptualising Consent Requirements' *Medical Law Review* 20, 191–226.

[17] Allen, C Joly, Y and Grandos Morena, P (2013). 'Data Sharing, Biobanks and Informed Consent: A Research Paradox? *McGill Journal of Law Health* 7, 85–120;.

[18] Kaye J et al 'Dynamic consent: a patient interface for twenty-first century research networks' *European Journal of Human Genetics* (2014) advance online publication 7 May 2014; doi: 10.1038/ejhg.2014.71 http://www.nature.com/ejhg/journal/vaop/ncurrent/full/ejhg201471a.html.

[19] Stein, D (2013) 'Reforming Biobank Consent Policy: A Necessary Move Away from Broad Consent Toward Dynamic Consent' *Genetic Testing and Molecular Biomarkers* 17 (12) 855–856; Steinbekk, K Myskja B and Solberg B (2013). 'Broad Consent versus Dynamic Consent in Biobanks Research: Is Passive Participation an Ethical Problem?' *European Journal of Human Genetics* 21, 897–902.

[20] Kush R and Goldman M (2014). 'Fostering Responsible Data Sharing through Standards', *New England Journal of Medicine* 370(23):2163–2165.

[21] National Cancer Institute, National Institutes of Health, US Department of Health and Human Services (2006). *First-Generation Guidelines for NCI-Supported BioRepositories*, http://biospecimens.cancer.gov/biorepositories/First%20Generation%20Guidelines%20042006.pdf.

[22] Sulston J and Ferry G (2003). *The Common Thread*, Corgi Books 165-9 and *passim*.

[23] Joly Y, Dove ES, Knoppers BM, Bobrow M & Chalmers D (2012). 'Data Sharing in the Post-genomic World: the Experience of the International Cancer Genome Consortium (ICGC) Data Access Compliance Office (DACO)', *PLoS Computer Biology* 8(7): e1002549. doi:10.1371/journal.pcbi.1002549.

[24] Caulfield T, Harmon SHE and Joly Y (2012). 'Open Science Versus Commercialization: a Modern Research Conflict', *Genome Medicine* 4:1.