Sideri Katerina and Graham Dutfield, (eds) 'Openness, Innovation and Science Policy in the Age of Data Driven Medicine' ***Special Issue Science and Public Policy*****.**

Introduction to the Special Issue, Graham Dutfield and Katerina Sideri.

Modern-day healthcare is becoming increasingly information intensive including at the personal level. The genomic data made available by the Human Genome Project gave this trend a great deal of impetus as did the more recent emergence of big data analytics alongside unprecedented computing power including artificial intelligence, enabling the generation of useful health-relevant information out of vast genomic and other datasets. Digitally recorded and annotated genetic and other molecular information acquired from large numbers of people, especially when combined with other information (lifestyle, family, personal electronic health history records etc.) coming from a wide range of sources, can provide not only a massive volume of health related data for analysis, but also diversity in kinds of information we can derive, from responsiveness to drugs, to likelihood of contracting particular diseases, and ways to prevent or reduce risk of certain diseases later in life. It is not just healthcare in the broad sense that is moving onto computer screens; medicine is becoming digital as much as it is chemical, especially when treatment concerns itself more and more with disease prediction, diagnosis, prognosis, and monitoring of sickness, health and treatment effects and side-effects, and of course with personalisation. According to one recent article on the subject, "the patient is an enormous repository of information that needs to be harvested as a partnership not only in clinical care but in discovery… The ability to stratify the phenotypic expression of wellness and disease will ultimately lead to better validation of human therapeutic targets for drug discovery" (D Ausiello, quoted in Elenco et al 2015).

Personalised medicine is one aspect, perhaps the most important, of the efforts currently being made by biomedical scientists and industry to enhance targeting of disease to achieve better health outcomes for more people. Personalised medicine deals with the tailoring of treatments in a way that responds to the variability of human beings, and to the fact that single diseases may really be families of sub-diseases. It implies individualisation of medical attention but whereas it does involve the testing of people for certain biomarkers conveying diagnostic or therapy-related information, those biomarkers are typically ones shared with other people. It is not a passing trend, but is reshaping the field of medicine.

Ideally, science policies would translate scientific research into technological innovation that benefits society. However, available medicines in many therapeutic areas are actually not very good even if they are very effective for some people. According to a recent study not one of the ten bestselling drugs in the United States helps the majority of patients who are given them. In fact, they benefit only 'between 1 in 25 and 1 in 4 of the people who take them. For some drugs, such as statins – routinely used to lower cholesterol – as few as 1 in 50 may benefit' (Schork 2015). In a 2001 article on pharmacogenetics, the authors found efficacy rates of major medicines in several areas to be very low: a 25 percent efficacy rate in oncology, 30 percent in Alzheimer's, 47 percent for hepatitis C virus, and 48 percent in osteoporosis to give a few examples (Spear, Heath-Chiozzi and Huff 2001). In addition, adverse drug reactions can cause deaths. Doubtless, some of these figures have improved in the years since then. For example, there are now medicines that can cure hepatitis C virus.

Thus, pharmaceutical companies' record of delivering truly innovative products in recent years is disappointing. In an influential article published in *Nature Reviews Drug Discovery*

(Scannell, Blanckley, Boldon and Warrington 2012),[1] the authors find that the decline actually began around 1950, whether or not it was perceived at the time. The authors posit that from that year onwards the number of medicines approved per billion dollars spent on research and development has on average halved every nine or so years. The fall is quite modest during the 1950s but then steepens from 1960. They call this phenomenon Eroom's Law (which is the famous and much more optimistic Moore's Law written backwards).

There is unlikely to be a single reason for this. However, the traditional innovation model based on the idea of strong intellectual property (IP) rights whose main function is to exclude competitors is seen by many critics as one factor. Given that data-driven medicine requires the mining of data of various types and from a wide range of sources, both public and private, weakness in the current IP-based innovation models seem likely if anything to become more serious. Why? Because IP rights inhibits sharing and collaboration at a time when these have never been more essential. Moreover, universities, patients and users participate in data driven innovation, which makes the picture more complicated. Moreover, exclusion and control sit uncomfortably with patients' and users' altruistic motives and universities' public mission. Unsurprisingly, the boundary between the open and the proprietary is passionately debated and constantly in flux: more open here, increasingly proprietary there, but with a tendency to be more of the latter and less of the former.

Openness in innovation is many times suggested as the solution. To some extent it is already being applied. As the Human Genome Project was coming to an end it became clearer than ever that much would be gained both scientifically and therapeutically from studying the genetic variability within the human species. One key unit of such variability was at the tiny level of the individual nucleotide base. Such variations shared by reasonably large numbers of people, and forming 90 percent of the genetic variability of our species, are called single nucleotide polymorphisms (SNPs). In 1999, a group of companies and research organisations together with the Wellcome Trust, then the world's largest medical charity, established the SNP Consortium. Its aim was to identify all of the common SNPs, of which there are now believed to be around 10 million, and map them onto the human genome. From the private sector, Glaxo Wellcome (as it was then called) took the initiative in starting such an endeavour but after meeting the Wellcome Trust and some other companies during 1998, it was decided that the ideal approach would be to establish a consortium. Funding came from the Trust and several large pharmaceutical company members and IBM and Motorola. In 2001 their shared data on SNPs was publicly released.

A similar pooling of public data was undertaken by the International HapMap Consortium, which comprised an international group of funders, government agencies and universities from the United States, the UK, Canada, China, Japan, Nigeria, as well as the SNP Consortium plus two biotech firms, Illumina and ParAllele Bioscience (The International HapMap Consortium, 2003). It has turned out that many SNPs throughout the genome are inherited together as 'blocks'. A haplotype is the arrangement of SNPs on each of these blocks. Given that the number of haplotypes is far lower than the quantity of SNPs, generating such a map offered an extremely convenient short cut in studying human genetic variability. The HapMap Project, the first phase of which was completed in 2005 (Goldstein and GL Cavalleri 2005), and the final one in 2009, required users to agree to a license that undertook them not to reduce access to the data or to pass data on to non-licensees.[2]

[1] The rest of this subsection draws on this article.
[2] For criticisms of the HapMap licensing policy, see Hope (2008), 308-9.

Such collaborative approaches and licensing schemes are not a rejection of intellectual property rights per se. Indeed, intellectual property protection may be necessary for 'open source' collaborative models to work. They do emphasize and seek to encourage collaboration of the kind that aggressive assertion of patent and other intellectual property claims would certainly preclude. The use of intellectual property rights is the best available sanction against those who acquire data and then seek legal protection covering elements of the received data and who may not be bound by any license. But of course the hope is that litigation is the last resort. This has been the case in software development, where open source collaborative models and licensing were first tried out with great success.[3] In fact, while the SNP Consortium's intellectual property procedures were for patent applications to be filed for any inventions arising, the point of doing so was not to claim monopoly protection but to record their priority dates so as to block patenting by others (Holden 2002). Currently the Structural Genomics Consortium, a charitable international partnership of non-profit foundations corporations has a policy of sharing freely and filing no patents on its discoveries.

Openness sound like a straightforward concept. However, its meaning remains elusive. On the one extreme we see such terms as "Open Data," "Open Software" and "Open Access". These present openness as a way to enhance transparency and collaboration (Benkler 2016), and preserve integrity and creativity.[4] On the other extreme, we see proprietary regimes, with "closed" data and legal rights such as patents being fundamental (the dominant pharmaceutical innovation model). However, commentators increasingly recognise that, for one thing, there are shades of openness. "Open innovation", a concept coined by Henry Chesbrough (2003), has been embraced by many in the pharmaceutical industry but while it does indeed involve sharing of knowledge, expertise and materials, it hardly deviates from the industry's patent-dependent business models. Even open source depends on copyright rules (Dusollier 2007). For another, the "closed" and "open" dynamically interweave (Hilgartner 2012). The latter relationship is the focus of the contributions of the special issue. Openness in this framework is not only a technical problem to be solved but also has a social, cultural, and moral facet.[5]

The question remains: how well do we strike the right balance in terms of promoting socially-optimal innovation? And how would one determine the ideal place to strike such a balance along the spectrum between closed innovation at one end and fully open at the other end? We hope this special issue will at the very least contribute to enhanced understanding of how these questions might best be resolved. Let us now turn to the contributions.

Regulatory rules and court decisions are one place where the balance between the open and closed is negotiated, but as we will see they are not the only 'spaces' where this happens. Recent decisions by the US Supreme Court placed genes and diagnostic methods outside the realm of patent protection proclaiming them non-eligible subject matter that cannot be the subject of private property. In Europe the balance continues to be struck in favour of more expanded eligibility in this area. On the other hand, the latter jurisdiction continues to have

[3] For an excellent and highly detailed discussion on the applicability of open source to biotechnology including reviews of several ongoing open source biotechnology initiatives, see Hope (2008).
[4] European Commission, (2014) "Consultation on 'Science 2.0': Science in Transition," Accessed November 21, 2017 http://ec.europa.eu/research/consultations/science-2.0/consultation_en.htm.
[5] Nuffield Council on Bioethics (2015) "The Collection, Linking and Use of Data in Biomedical Research and Health Care: Ethical Issues." Accessed November 25, 2017 http://nuffieldbioethics.org/wp-content/uploads/Biological_and_health_data_web.pdf; Michael A. Peters. "Open Science, Philosophy and Peer Review." *Educational Philosophy and Theory* (2014) 46 (3): 215–19.

broader exclusions in other areas of biomedicine as well as a ban on the patentability of inventions whose commercial use would be immoral and contrary to *ordre public*. However, as our contributors discuss, extending legal rights to one type of scientific research product that we preclude from others affects how societies innovate. Shubha Ghosh explains that denying patents may influence the shift of commercial activity from genes to genetic data mining. In this sense, court decisions influence politics and the future development of industry in indirect ways. Katerina Sideri argues that these decisions reflect policy choices and a particular understanding of the proper role of the state in regulating the marketplace and knowledge production in the emerging information economy. In their respective contributions, both authors claim that these decisions essentially endorse practices of data processing which constitute a new type of public domain necessary for fueling the development of the new data-mining and analytics tools and the next generation of data intensive therapeutics in the field of data driven medicine. These technologies learn from data to predict the future behavior of individuals in order to drive better decisions.

Outside of the courts, and perhaps more significantly, the relationship between 'open' and 'closed' in innovation depends on social norms and values (Botsrom, 2017). In this sense, openness often invites the notion of participation, which is infused with the ideas of social solidarity and altruism, a far cry from the idea of impersonal exchanges in vast global markets that property rights imply. Thus, the social meaning of openness is negotiated in the context of clinical applications of revolutionary technologies such as cell-free foetal DNA prenatal testing, which simplifies testing for abnormalities in the foetus. Naomi Hawkins discusses the rapid development and widespread adoption of the technology in the clinic around the world, and employs qualitative analysis of interview material with users of technologies to question the extent to which and the reasons for failure to comply with patent law. This approach resonates well with Shobita Parthasarathy's contribution advocating the expansion of qualitative research on patents and intellectual property related to innovation. There is urgent need for policy makers to go beyond economic analysis so as to come to grips with the broader implications of intellectual property for social and political orders. In fact, Shobita Parthasarathy argues that this qualitative research can help governments produce patent decisions and policies that are both more socially beneficial and politically legitimate.

To turn to other legal and policy developments in Europe that go beyond patents, Timo Minssen, Rajam Neethu and Marcel Bogers discuss openness in the context of initiatives with respect to transparency of clinical trial data and note the potential tensions with the General Data Protection Regulation (GDPR) in the EU. For the GDPR the question of openness is viewed through the angle of user control of data in the era of big data analytics, but the policy goals behind transparency of clinical trial data seem to be quite different: the focus is on promoting science and open innovation that will benefit society at large and is based on the understanding that data sharing and open innovation go hand-in-glove. Highligting the tensions between these two levels of openness is imporatnt and links to the more theoretical discussion of Barbara Prainsack's paper, the last contributor of the special issue. Barbara Prainsack elegantly summarizes the theoretical nuances of the notion of openness in the context of Personalised and Precision Medicine. Her argument is that different theorizations pose different goals for public policy. She identifies three ways to theorize openness: the 'ontological,' the 'pluralistic' and the 'emancipatory'. The 'ontological sense' relates to openness at the level of the person, the 'pluralistic sense' brings to the foreground the plurality of perspectives and values, while the 'emancipatory sense,' poses questions with regard to concentrations of power of corporate actors dominating the field of innovation in data driven

medicine. Depending on the choice of theoretical approach the focus of key goals for public policy shifts.

In short, the special issue seeks to discuss the notion of openness in data driven medicine and contributors are social scientists who contribute to this debate by looking into the questions economists cannot answer. All contributors write in the fields of law, political science and Science and Technology Studies.

References

Dusollier, S. (2007) 'Sharing Access to Intellectual Property Through Private Ordering' *Chicago-Kent Law Review,* 82/3: 1391-1431.

Henry W. Chesbrough (2003), *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Boston: Harvard Business School Press.

Benkler, Y. (2016) 'Open Access and Information Commons,' in Francesco Parisi, ed. *Oxford Handbook of Law and Economics: Private and Commercial Law* (Oxford: Oxford University Press).

Bostrom, N. (2017) 'Strategic Implications of Openness in AI Development,' *Global Policy* 8/2: 135–148.

Elenco, E., L. Underwood and D. Zohar (2015) 'Defining Digital Medicine', Nature Biotechnology, 33(5): 460-1.

Goldstein, D. B. and G. L. Cavalleri (2005) 'Understanding Human Diversity', Nature, 437: 1241-2.

Hilgartner, S. 'Novel Constitutions? (2012) 'New Regimes of Openness in Synthetic Biology,' *BioSocieties* 7/2: 188–207.

Hilgartner, S., (2012) 'Selective Flows of Knowledge in Technoscientific Interaction: Information Control in Genome Research,' *The British Journal for the History of Science* 45/2: 267–80.

Holden, A. L., (2002) 'The SNP Consortium: Summary of a Private Consortium Effort to Develop an Applied Map of the Human Genome', Biotechniques, 32: S22-S26.

Hope, Janet (2008) *Biobazaar: The Open Source Revolution and Biotechnology*. Cambridge: Harvard University Press.

The International HapMap Consortium (2003) 'The International HapMap Project', Nature, 426: 789-96.

Scannell, J. W., A. Blanckley, H. Boldon and B. Warrington (2012) 'Diagnosing the Decline in Pharmaceutical R&D Efficiency, Nature Reviews Drug Discovery, 11: 191-200.

Schork, N. (2015) 'Personalized Medicine: Time for One-person Trials, *Nature* 520: 609-11.

Spear, B.B., M. Heath-Chiozzi and J. Huff (2001) 'Clinical Application of Pharmacogenetics', Trends in Molecular Medicine, 7(5): 201-4.