Title	Moving beyond Consent for Citizen Science in Big Data Health Research
Author(s)	Cheung, ASY
Citation	Cheung, Anne S. Y., Moving beyond Consent for Citizen Science in Big Data Health Research (March 29, 2017). Available at SSRN: https://ssrn.com/abstract=2943185 or http://dx.doi.org/10.2139/ssrn.2943185
Issued Date	2017
URL	http://hdl.handle.net/10722/241269
Rights	This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Moving Beyond Consent for Citizen Science in Big Data Health Research

Anne SY Cheung

Introduction

Consent has been the cornerstone of personal data privacy regime.[1] It authorizes the collection, use and process of personal data. When it comes to health and medical research, consent is a prerequisite for the intervention to one's body, the collection of bio-specimens and the use of personal data. The doctrine of consent is premised on the liberal tenets of individual autonomy, dignity and integrity, rooted in the fundamental respect to a person and intertwined with the right to respect for privacy.[2] More important, consent is only meaningful if it is freely given (voluntary), specific and informed (Article 29 Data Protection Working Party 2011, 29).

The above concern is particularly pertinent to citizen science in health and medical research, in which the nature of research is often data intensive with serious implications for individual's privacy and other interests. Although there is no standard definition for citizen science, the European Commission has highlighted its general features to be the gathering and volunteering of data by non-professionals, and the participation of non-experts in analysis and scientific experimentation, with public input into research and project.[3] Citizens become experimenters, stakeholders, purveyors of data, research participants or even partners in the process.[4] Consent from

[1] Organization for Economic Co-operation and Development, OECD Privacy Framework 2013 at http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf.
[2] (Beauchamp and Childress 2012, 107)
[3] (European Group on Ethics in Science and New Technologies to the European Commission 2015, 23).
[4] By citizens as experimenters, this refers to patients participating in various degrees in experimentation; as stakeholders refers to patient expert groups; as purveyors of data refers to citizens or patients sending data through digital devices, mobile devices and other information communication

citizen scientists is indispensable as it is a constitutive element for self-determination and self-empowerment for participants. Furthermore, consent from participants as data subjects determines the responsibility and accountability of data users. Under orthodox understanding, consent should be given on a one-on-one basis, for a "single study at a single institution for a specific purpose"(Kuehn 2013).

While all these may sound sensible and reasonable, technology always produces new directions and challenges for research. With the advancement in information technology and data analytics, bio-medical and health research have become data-intensive, global and virtual (Schmietow 2016, 197). Bio-banks and virtual research repositories are gaining prominence and significance (Auffray et al. 2016). At the same time, the risk inherent in health science research and big data technology has often extended beyond the existing data. In addition, the use and transfer of data for other unforeseen purposes is often outside the control of the original research team. Plus, risks and harm of subsequent data use may not be known at the time of data collection. Hence, consent becomes problematic because the traditional standard and understanding on consent that one can fully specify the terms of agreement in advance becomes questionable in big data science and citizen science.

We cannot help but ask: is consent still valid? Should it still be one of the critical criteria in citizen science health research which is collaborative and contributory by nature? While the big data challenge is not unique to citizen science health and medical research, the inherent sensitivity of health data exacerbates the problem which calls for close scrutiny of the doctrine of consent. With a focus on citizen science health and medical data research, this study examines the doctrine of consent and its inadequacies. It then analyzes the alternative and adaptive models of

technology. (European Group on Ethics in Science and New Technologies to the European Commission 2015, 25).

"open consent" [5] and "dynamic consent"(Kaye et al. 2015). Facing the challenges that big data and citizen science pose to personal data protection and privacy, this article explores the legal, social and ethical concerns behind consent. It argues that navigating one's way through different models of consent and the varied choices in consent forms can be a legal minefield. We need to move beyond the consent paradigm into a broader framework of accountability, taking into consideration of harm and risk assessment. Ultimately, what lies behind consent are the entailing values of autonomy, fairness and propriety in the name of research.

I. Citizen Science and Big Data Health Research

We begin our discussion by looking at the nature of citizen science. In the mid-1990s, Alan Irwin popularized and prompted the discussion on citizen science (Prainsack 2014). By which, Irwin refers to the examination of the relationship between citizens and science, including science that assists citizens' needs and concerns, and science developed and enacted by citizens themselves (Irwin 1995, xi). Irwin flags up the need to face up the challenges posed by risks in scientific research and sustainable development.[6]

By the 21st century, citizen science has developed into different forms of participation by non-professional scientists, exhibiting various dimensions in the co-operation between professionals and non-professionals, opening up multiple levels of engagements in the area of health sector. At one end of the spectrum, there is citizen led, patient owned initiatives of sharing quantitative information, exchanging experience on treatment and searching for the right clinical trials on online platforms.

[5] (Angrist 2009).
[6] Irwin's case study was mainly on environmental development.

Prominent examples include PatientslikeMe,[7] CureLauncher[8] and CureTogether.[9] At the other end of the spectrum, there is commercial or government led research. For instance, Sage Bionetworks took advantage of smartphone based health technology to study the lifestyle of 17,000 Parkinson's disease patients in 2015.[10] It also paired up with Apple ResearchKit to study the life quality of 60,000 breast cancer survivors.[11] In the same year, US President Obama announced the nationwide US$215 million Precision Medicine Initiative to build a large-scale research enterprise between public and private sectors, calling for one million volunteers to contribute their health data so as to extend precision medicine to all diseases (The White House 2015). In between the two models, there is a third one catering for joint collaboration between citizens and health professionals in creating knowledge. Prominent example includes the Sarroch Bioteca Foundation founded in 2012 in pursuant of a "citizen veillance on health" project in Italy (Tallacchini, Boucher, and Nascimento 2014). The project was launched by the Sarroch municipality in 2006 to gather biological samples donated by citizens so as to monitor genetic changes as health indicators in relation to the environment. All citizens of the municipality were invited to become members of the project. The aim was to use science for policy in the regulation health and institutional implementation. The Sarroch example was seen as a joint effort for collective governance and a model for democratic health choice.[12]

Regardless of the level of citizen involvement, the degree of participation, or the form of co-operation in the above models as contractual, contributory,

[7] https://www.patientslikeme.com
[8] http://curelauncher.com
[9] http://curetogether.com
[10] (European Group on Ethics in Science and New Technologies to the European Commission 2015).
[11] (European Group on Ethics in Science and New Technologies to the European Commission 2015).
[12] (European Group on Ethics in Science and New Technologies to the European Commission 2015, 28).

collaborative, co-creative or collegial,[13] data privacy of research subjects and security issues will be triggered whenever citizens have contributed their data or biospecimens to the projects. Following the contemporary new wave of citizen science research in big data is a whole new set of legal and ethical concerns. First, we are witnessing unprecedented scale of online crowd sourcing, with researchers pooling data together using big data capture strategies and data analytics (Rothstein, Wilbanks, and Brothers 2015). As progress of research often extends beyond the existing data, big data technology use and transfer of data for other unforeseen purposes is often outside the control of the original research team. Second, despite promise of data de-identification, third parties can match data sets to re-identify individuals (Sweeney 1997). Third, with the advancement of data mining and big data technologies, risks and harm of subsequent data use may not be known at the time of data collection and use. Namely, researchers of existing team or even third parties can match data sets to re-identify individuals. For instance, large scale harvesting of health data can reveal unnoticed correlations between lifestyle and medical conditions of individuals, which are important information to insurance companies ("Tech Companies See Market Opportunity in Healthcare Innovation" 2016). In other words, consent becomes problematic in citizen science research especially in the big data era. The model that one can fully disclose and specify the terms of notice and consent at the outset has become illusory.

[13] The typology of the five project models is formulated by Shirk and her colleagues. Contractual projects refer to communities asking professional researchers to conduct a specific investigation and report; contributory projects refer to those designed by scientists, with citizens contributing data; collaborative projects is similar to contributory project except with citizens helping to refine project design, analyze data or disseminate findings; co-created projects are jointly designed by scientists and citizens, in which some citizens are actively involved in most or all aspects of the research process; and collegial contributions refer to non-credential individuals conducting research independently with varying degrees of expected recognition by the professional scientists. (Shirk et al. 2012)

II. Rethinking Consent

Consent has been a cardinal doctrine in clinical treatment and research. It is premised on the respect for individual autonomy which embodies the principle of self-rule that is free from "controlling interference by others and limitations that prevent meaningful choice"(Beauchamp and Childress 2012, 101). It is enshrined in numerous international treaties, legal guidelines and code.[14] Namely, on the protection of human rights, the International Covenant Civil and Political Rights (ICCPR) stipulates that "free consent" is a prerequisite for medical and scientific experimentation.[15] On personal data, the General Data Protection Regulations (GDRP) of the European Union states that "explicit consent" is necessary for the processing of genetic, biometric and health data.[16] On experiments done by physicians, the Nuremberg Code sets the standards to which physicians must conform when carrying out experiments on human subjects which requires the obtaining of consent and the ascertainment of competence from human subject in experiments.[17] On medical research, the Declaration of Helsinki of the World Medical Association calls for "informed consent, preferably in written form" from physicians.[18] Again, on bio-

[14] For an historical overview, see (Meier 2002) For further comparison of global guidelines on consent and informed consent, see (Bhutta 2004)

[15] Article 7 of the ICCPR "No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment. In particular, no one shall be subjected without his free consent to medical or scientific experimentation." (ICCPR 1976)

[16] Article 9 of the General Data Protection Regulations (GDPR). In addition to article 9, article 7 of the GDPR requires that the request for data subject's consent must be clearly distinguishable, intelligible, easily accessible, and expressed in clear and plain language. The GDPR will come into force in 2018 (Articles 94 and 99 of the GDPR). Regulation 2016/679 of the European Parliament and of the Council, General Data Protection Regulation, Official Journal of the European Union, L119/1, 27 April 2016 at http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf.

[17] Article 1 of the Nuremberg Code states that " the voluntary consent of the human subject is absolutely essential . . . [and includes] legal capacity . . . free power of choice . . . sufficient knowledge and comprehension of the [nature, duration, and purpose of the experiment] . . . to make an understanding and enlightened decision." The Nuremberg Code (1947) at http://www.cirp.org/library/ethics/nuremberg/.

[18] WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects Promulgated in 1964 and revised eight times since. Article 26 states that the physician should obtain

medicine study, the Convention on Human Rights and Biomedicine calls for "free and informed" consent; [19] and the International Ethical Guidelines for Biomedical Research Involving Human Subjects insists that investigators must obtain the "voluntary informed consent of the prospective subject …and that waiver of informed consent is to be regarded as uncommon and exceptional and must in all cases be approved by an ethical review committee."[20] On research involving human subjects in general, the US Code of Federal Regulation specifies that informed consent from participants in research must involve discussion of the nature of the involved procedure, its risks and benefits, and alternative treatments available. [21] This then must be concluded with the free assent from participants.[22]

Regardless whether it is free, explicit, informed and voluntary, the four essential elements of a valid consent are comprehension or understanding, voluntary participation, competence and disclosure (Sreenivasan 2003). While the first three refer to the duty of doctors or researchers to obtain the voluntary agreement of human subjects before participation, the last one refers to their duty to disclose adequate information to the subjects.[23] Both limbs are integrated requirements of consent as a single legal and moral doctrine.[24] Seemingly, the above frameworks in international and national law, codes and guidelines have provided the necessary and sufficient legal basis for informed consent and for the use of one's data or bio-specimens. But

---

the "subject's freely-given informed consent, preferably in writing" at http://www.wma.net/en/30publications/10policies/b3/

[19] Council of Europe, Article 5 of the (Council of Europe 1997)

[20] Guideline 4 of the Council for International Organizations of Medical Science (2002) published in 1993, and since revised. Investigators must obtain the "voluntary, informed consent" of the prospective subject [or legally authorized representative]. . . . Waiver of informed consent is to be regarded as uncommon and exceptional, and must in all cases be approved by an ethical review committee." at http://www.cioms.ch/publications/guidelines/guidelines_nov_2002_blurb.htm

[21] §46.116 Code of Federal Regulations, TITLE 45 PUBLIC WELFARE DEPARTMENT OF HEALTH AND HUMAN SERVICES PART 46 PROTECTION OF HUMAN SUBJECTS (Revised January 15, 2009, Effective July 14, 2009).

[22] §46.116 Code of Federal Regulations.

[23] §46.116 Code of Federal Regulations.

[24] §46.116 Code of Federal Regulations.

how detailed should description of the research be and how much disclosure would be required as adequate remains controversial.

In the context of data-driven citizen science research, participants have to face the additional uncertainty and unpredictability of research progress. For example, in the 1980s, a group of Canavan-disease affected families developed a disease registry and tissue bank to encourage research in the area.[25] They provided tissue for research on the disease and aided in the identification of other affected families.[26] With three nonprofit organizations, they developed a confidential database and Canavan disease registry, which had attracted financial sponsorship. However, when one of the chosen physician-researcher decided to isolate and patent the Canavan gene sequence and developed genetic screening tests for it, the families sued the researcher and his institution.[27] The bitter legal battle ended only in half victory. The U.S. District Court for the Southern District of Florida dismissed several of the plaintiffs' claims, including lack of informed consent, breach of fiduciary duty, fraudulent concealment of the patent, and misappropriation of trade secrets. Nevertheless, the court upheld the claim of unjust enrichment made by the donors of tissue, on the grounds that "the facts paint a picture of a continuing research collaboration that involved Plaintiffs also investing time and significant resources."[28]

In more recent times, another notorious example is the 23andMe project. It started out in late 2007 as a company offering genetic testing at a very low price of US$299, giving out spit-kits and asking for saliva samples (Seife 2016). The way of

[25] Canavan disease is a progressive, fatal neurological disorder that begins in infancy. It is caused by an inherited genetic abnormality resulting in improper transmission of nerve signals. Canavan Foundation, Canavan Disease, http://www.canavanfoundation.org/about_canavan_disease. For discussion of the dispute, see (Evans 2016).
[26] Greenberg v. Miami Children's Hospital Research Institute.264 F. Supp. 2d 1064.
[27] Greenberg v. Miami Children's Hospital Research Institute.264 F. Supp. 2d 1064.
[28] Greenberg v. Miami Children's Hospital Research Institute.264 F. Supp. 2d 1064. .

testing was perceived as a "fun way" to learn about one's genetics.[29] In 2008, 23andMe added a new research feature named 23andWe which played up citizen science rhetoric, emphasizing on strong participatory features.[30] It invited customers and participants to vote for disease that the company promised that they would then prioritize in their research.[31] In turn, the customers and participants were asked to contribute their lifestyle and other relevant information for research purpose. By that time, the company had offered free spit kits to people who had been diagnosed with the type of disease that the company would like to focus and to do research.[32] The price of spit-kit had also been lowered to US$99.[33] By 2012, 23andMe had about 15,000 users.[34] The mode of business operation was a commercial company that drew heavily on the participation of citizen scientist. Patients and members of patient support groups joined under the impression that they were contributing their genetic and personal data for the development of treatment and long term research. However, they were soon to wake up to reality when the company filed a number of patent applications later in 2012.[35] People also came to realize that 23andMe was sharing aggregate data about its customers and participants to third parties, and Google has been investing in the company.[36] Consequently, the company was under severe criticism but defended that it had informed the customers and participants all along in their Terms of Services and consent forms.[37] Although what 23andMe had done was technically lawful, Barbara Prainsack pointed out that it was dishonest and immoral

[29] Greenberg v. Miami Children's Hospital Research Institute.264 F. Supp. 2d 1064.
[30] (Prainsack 2014, 156).
[31] (Prainsack 2014, 156).
[32] (Prainsack 2014, 156).
[33] (Seife 2016).
[34] (Prainsack 2014, 156).
[35] (Prainsack 2014).
[36] In fact, the founder of the company, Anne Wojcicki, was the wife of Google-boss Sergey Brin at that time. (Prainsack 2014, 158); (Seife 2016).
[37] (Prainsack 2014).

for the company to capitalize on the "free labour" and data capital of its customers, patients and participants for profit earning under the grand name of research.[38] In particular, Prainsack observed that the business model of citizen science in 23andMe was "continually evolving."[39] It was highly unlikely that the participants could keep up with the frequent modifications of the Terms of Service and have read the fast changing, constantly updated terms in small print on the website.[40]

What the above two incidents have illustrated is that there is misalignment of orientations between citizen science participants and expert researchers. The former group was motivated by a genuine commitment to facilitate disease research, to contribute to health knowledge and to create collective benefits. In contrast, the latter was motivated by profit earning, and individual or corporate success. Yet, this mis-match might not be present at the outset of research projects. Rather, it is due to the fluid and flexible nature of health and medical data research and citizen science that the projects soon spin out of control of the citizen science participants and evolve beyond their own meanings.

In addition to the problem of expert researchers gradually growing apart from their research subjects, one may also notice that the relation between expert researchers and participants is not on a traditional model of one-to-one basis. Instead, it is rested on an elaborate network backed by complex organization structure, and staffed by different experts at various levels. Citizen scientists or participants motivated by altruism to share their personal data mistakenly thought they could retain some form of control in a collaborative or cooperative manner.[41] Their

[38] (Prainsack 2014).
[39] (Prainsack 2014, 157).
[40] (Prainsack 2014).
[41] (European Group on Ethics in Science and New Technologies to the European Commission 2015, 52).

solidarity, sadly, was later exploited by researchers or commercial groups in both public and private spheres.

Furthermore, if informed consent requires disclosure by researchers and comprehension by participants, full disclosure of information will become neither definable nor achievable at the outset of the research due to the fast changing nature of research which is heavily dependent on data analytics. Writing on clinical treatment and research, Onora O'Neil has remarked on the inherent deficiency of informed consent as a doctrine (O'Neil 2002, 42–43). She explains that this is not due to any deficiencies in the procedures to ensure informed consent has been fulfilled, but more to the fact that consent is a "propositional attitude."[42] It is a "*description* of a *proposal*" for treatment or research.[43] One can only consent to the specific descriptions of a proposition but may not be aware of the foreseeable consequences, or those entailed to it.

Does this then mean that informed consent is no longer valid? Alternatively, does it suggest it is high time for an urgent refinement of the requirement of consent? Mittelstadt and Foridi have argued that the traditional framework on informed consent "does not clearly transfer" to research involving biomedical big data (Mittelstadt and Floridi 2016, 311). They point out this is because the doctrine of informed consent is formulated for single, specific research or treatment but not for the sharing, aggregating or repurposing of data that may reveal unforeseen information.[44] As a result, until we have found a satisfactory alternative model, the pressing concerns on obtaining informed consent for citizen science research in health data remains to be: deciding the purposes the data are collected and the duration that the data will be kept;

[42] (O'Neil 2002, 43).
[43] (O'Neil 2002, 43), emphasis by O'Neil.
[44] (Mittelstadt and Floridi 2016, 312).

identifying who is permitted to have access to the data and who is processing the data for what purposes; and determining what to do in case of misuse of data.[45]

III. Re-negotiating Consent

To tackle the above, different models on informed consent have been formulated by researchers in this area. Here, we evaluate the common forms of open and dynamic consent, including variations of the latter known as portable and meta consent.

1. Open Consent

In light of the inherent uncertainty and unpredictability of big data health research, some scholars have advocated for veracity or "radical honesty"(Wilbanks 2014) in the model of "open consent," which deliberately excludes any promises about privacy and requires participants to demonstrate comprehension of the nature of the research and the risks involved prior to enrollment.[46]

This has been used in the famous Personal Genome Project (PGP) by Harvard University since 2005.[47] The aim of the project is to test DNA sequencing technologies on human subjects by building a database of human genomes and traits, with ambition to be a global network project.[48] The nature of the database is open source, open access, and participatory and collaborative. The target is to collect the

[45] (European Group on Ethics in Science and New Technologies to the European Commission 2015, 53).
[46] (Ball et al. 2014). Other examples of open consent model include the Omics project, The Human Microbiome Project and the American Gut Project. See (European Group on Ethics in Science and New Technologies to the European Commission 2015, 14–16).
[47] http://www.personalgenomes.org/.
[48] At the time of writing, UK, Canada and Austria have joined the network.

genomes of 100,000 individuals and to make the information public with no serious effort at deidentifcation.[49] Since DNA is the ultimate digital identifier of an individual but de-identification of samples would impoverish the data,[50] the PGP research team has decided to be forthright and honest with the participants, aiming that they should be "truly informed" about the nature of the research. Participation by public is encouraged and volunteers are asked to give "open consent" to ensure that they understood the scientific nature of the experiment and that they also understood that privacy and confidentiality could not be guaranteed.

Misha Angrist, being one of the original ten participants of the PGP back in 2006, shared his experience and reflections. According to him, participants had to go through an eligibility screening process first which included filling out a questionnaire regarding family circumstances and privacy preferences.[51] Second, they would review a study guide that covered the potential risks of participating. Third, they then took an "entrance exam" that covered the areas of how PGP worked, knowledge of genetics, ethical principles governing human subjects research; and their comfort level with having their genome and health records in the public domain. They had to score 100% in the exam before they could be enrolled in the project. Finally, they had to sign a consent form, which stated the possibility of reidentification, disclosure of nonpaternity and loss of insurance, among other risks of embarrassment, discrimination, data loss and any unforeseen problems.

Participants were described as "co-drivers" of the project (Angrist 2009). They were expected to have solid knowledge about the field. Take Angrist as an example, he himself is a scientist working for an established institute for genome

[49] (Angrist 2009).
[50] (Angrist 2009).
[51] (Angrist 2009).

science and policy.[52] He completed an early version of the entrance exam and suggested changes to certain questions, and questioned the rationale for specific analyses(Angrist 2009). He was also one of the three initial ten participants who served on the PGP Board of Directors(Angrist 2009). He was careful enough to carry out certain test on his genotypes and make sure the result was negative before deciding to make his cell line available to the public (Angrist 2009). However, it is doubtful how many other citizen scientists or participants could achieve such thorough understanding of the research and its implications to privacy.

Other than the fact that the nature of research and open consent require advanced level of comprehension from citizen science participants, scholars have criticized that open consent is far from "true consent"(Caulfield, Upshur, and Daar 2003). First, open consent does not allow participants to act meaningfully on their continuing interest in their own health data.[53] Second, the way that it operates does not take prevention of harm or controlling use into account. Third, equally worse, the model is a distortion of informed consent as it does not include permission for unforeseen research, re-contact of subjects, the right to withdrawal, the setting of time limits on the use of data, the restriction of information or materials to third parties, and information on implications for groups and information on commercial uses.[54] Altogether, so called veracity has become an excuse to absolve researchers from accountability and responsibility. Although open consent maybe legally valid, its practice remains ethically vague and questionable.

2. Dynamic Consent

[52] Affiliated with Duke University Institute for Genome Sciences & Policy. Misha Angrist at https://sanford.duke.edu/people/faculty/angrist-misha.
[53] (Angrist 2009).
[54] (Greely 2001).

Rather than asking participants to take a leap of faith into uncertainty, a slightly refined model of "broad consent" has been proposed. While one gives consent to a framework for future research of certain types, ethical review of each specific research project by an independent ethics committee is required (Steinsbekk, Kåre Myskja, and Solberg 2013, 897). In addition, researchers must provide strategies on how to update regularly the participants and how to enable ongoing withdrawal opportunities for the participants. Examples of broad consent model research are the UK Biobank project and the Norwegian Mother Child Cohort Study.[55] Nevertheless, regular updates to participants in ongoing research is seen as "extras" in this model (Kaye et al. 2015, 142). Thus, legal and ethical concerns have been raised as to whether broad consent is a form of genuine informed consent when participants are reduced to be passive subjects rather than research partners (Teare et al. 2015, 2).

Approaching consent from a different perspective finally comes the alternative model of "dynamic consent." It is a model tailor-made to the need of participants by utilizing online interface and information technology based platform. Information about specific use of personal data and tissue and requests for consent for such use will be put to the participants through online platform (Kaye et al. 2015). Participants are allowed to engage in an interactive personalized interface as much or as little as they choose and to alter their consent choices in real time (Kaye et al. 2015, 142). Consent is seen as a process, an ongoing interaction between researchers and participants. Hence, consent becomes dynamic because it allows participants to interact with the researchers over time, to consent to new projects, to alter their consent choices in light of any new circumstances. This model was first designed for

[55] (Kaye et al. 2015).

the EnCoRe project of three biobanks in Oxford (2008 to 2012).[56] Another example is the Registries for All (Reg4All) project run by Genetic Alliance in partnership with the technology company Private Access (Mathews and Jamal 2014). Reg4All allows participants to decide how their data are being used and shared with particular researchers, institutions, or people studying a specific disease. They will also be able to track who have used their data and how.

Scholars praise the model of dynamic consent as providing "personalized communication interface for interacting with patients, participants and citizens"(Kaye et al. 2015, 142), implementing engagement 2.0 in the era of Web 2.0(Teare et al. 2015, 1). Apparently, it enables consent to be given to multiple researchers and projects, to open-ended and ongoing research, and to the use of secondary research or down-streaming of data-use. Besides, dynamic consent has overcome the problem of locked in consent confined to one experimental procedure for granting autonomy, choice and control to individuals. In other words, participants can electronically control consent through time and receive information about the uses of their data(Williams et al. 2015). At the same time, researchers can also manage the necessity to re-contact and to seek "re-consent" participants much more easily.

Understandably, dynamic consent as a participant-centric initiative has its special appeal. Refining the model of dynamic consent, there are further variations to it.

(a) Portable Legal Consent (PLC)

Another model is "portable legal consent" proposed by John Wilbanks of Sage Bionetworks (Wilbanks 2014). Under this model, individuals are recognized that they have rights with respect to the data generated from their bodies. They therefore will

[56] (Kaye et al. 2015, 145.) (Mont, Sharma, and Pearson 2012).

decide the kind of data that they would like to donate and share. For example, in the Sage Bionetworks, the suggested five categories are genetic sequence, clinical information, medical record, patient reported outcomes and personal sensor data.[57] Consent is not tied to any particular study but carried around by the participants like organ donation status (Kuehn 2013). In that sense, consent becomes portable and controllable. Obtaining consent will be done through an online interactive consent system. Participants can share their own data broadly in the public domain to serve scientific research regardless of the particular institution involved. They can easily change their consent status for future study or secondary use of data. In turn, the database of genomic information being collected through portable legal consent will be available to anyone who agrees to its terms. These include not to use the data to harm anyone or to identify the participants.[58] Users also agree to publish their work based on open access policy.

(b) Meta Consent

Rather than focusing on the distinct categories of personal data, the meta consent model allows participants to express a preference for how and when to provide consent at a metal level, i.e. how and when they would like to be presented with a request for consent to the use of their personal health data and biological material (Ploug and Holm 2016, 724). In this model proposed by Ploug and Holm, participants must be provided with a predefined set of types of consent;[59] a set of types of data;[60] and a set of types of research contexts to choose from.[61]

[57] (Wilbanks 2014, 248).
[58]Sage Bionetworks, https://s3.amazonaws.com/static.synapse.org/governance/SageBionetworksSynapseTermsandConditionsofUse.pdf?v=4
[59] Ploug and Hom mention three types of consent. The first is specific consent referring to consent request for each new specific project using data, but not for each and every use of data.

While acknowledging similarities with dynamic consent, Ploug and Holm argue that meta consent is different in the sense that the model of dynamic consent is originally designed for bio-banks(Ploug and Holm 2016, n. 29). In contrast, meta consent is developed with the aim to handle and configure consent preferences for the entire population for all kinds of data and biological samples, with a vision that every citizen is a potential participant in big data research especially in medical research(Ploug and Holm 2016, n. 29). The meta consent model is designed "to provide a definitive answer by letting individuals design future consent requests on the basis of predefined types of consent, data and contexts"(Ploug and Holm 2016, n. 29).

3. Consent and its Limitation

Regardless which variation or refinement of dynamic consent that one chooses, problems remain. First, dynamic consent is in essence an information governance model, which is useful only for the well informed, engaged and e-health literate participants (Williams et al. 2015). Concerns of digital divide and social exclusion have yet to be addressed.[62] Second, participants will be asked for consent continuously because each new project requires fresh new consent to be given. Arguably a person may potentially receive hundreds of consent requests each year.[63] This is likely to cause routinization of consent behavior resulting in people not

---

The second is broad consent for "broader categories of research." The third type is blanket consent and blanket refusal for one-off decisions concerning participation or non-participation in research. (Ploug and Holm 2016, 726).

[60] This includes data from electronic patient record, tissue/genomic data, health databases and linkage to non-health data (Ploug and Holm 2016) .

[61] Research context refers to private versus public, commercial versus non-commercial, and national vs international (Ploug and Holm 2016).

[62] (Williams et al. 2015) (Steinsbekk, Kåre Myskja, and Solberg 2013, 898)

[63] Ploug and Holm discusses the hypothetical experience of a resident of Denmark if specific consent is required for every secondary use of data. (Ploug and Holm 2016, 723)

reading the information and not reflecting the choice but simply choosing habitually to consent or refuse to give consent (Ploug and Holm 2016, 724). Although the refined model of meta consent allows opt-out or broad consent, it is skeptical whether such is allowed under the new European Union regime of GDPR, and whether it is considered a form of valid informed consent. Third, the dynamic consent model has seemingly passed control and choice to individual participants. Yet at the same time, it has also passed responsibility to them without ensuring that they have the required knowledge and competence to make informed decisions (Steinsbekk, Kåre Myskja, and Solberg 2013, 900). Steinsbekk et al have even argued that based on the above point, broad consent may be better than dynamic consent as independent review from research ethics committee is required in the former (Steinsbekk, Kåre Myskja, and Solberg 2013, 900).

Overall, the open consent model and the dynamic consent model may have enabled better participation of participants but they remain largely information strategies. The mere passing of more information to participants, and the seeking of their indication at different stages of research do not necessarily amount to building a democratic and participatory model of medical research. Steinsbeek et al point out that the participation envisaged is limited as it is "participation *inside* an already established research arena where only minor changes of policy are up for discussion"(Steinsbekk, Kåre Myskja, and Solberg 2013, 900). Furthermore, at best, we have filled only part of the knowledge gap (mentioned at the end of Part II of this article) in enabling participants to find out more about the purposes of data collection, the persons who have been accessing their data and "empowering" them to have more control on how the data is being used down the stream of data re-use. The success of the two alternative models is dependent highly on how informed, competent,

knowledgeable and reflective of the participants to engage meaningfully. However, we have not addressed the nightmare scenario of what to do in case things go wrong.

IV. Beyond Consent: The Model of Accountability

Indisputably, consent plays an important role in health data research, but mere information disclosure and seeking of participants' indication of choices do not guarantee respect and self-determination of individuals. At most, the above suggested alternative consent frameworks have fulfilled the contractual ritual required in law (Felt et al. 2009, 101). They may have "managed" the legal concerns (Felt et al. 2009, 102), but they have not resolved the problems of risks and harms not mentioned in the terms of agreement. Rather than shifting across different modes of consent and putting participants through strenuous, targeted exercise of choices and forms, the consent model should be complemented with an accountability model.

Big data technology has opened up undreamed of capacities to gain sophisticated understanding about the way that we can process and use data so as to organize our society and our lives. Those insights, unfortunately, can be pitfalls at the same time. Governments at different jurisdictions are eager to capture the benefits of big data but to weed out its harm (van der Sloot and van Schendel 2016, 116–17). In the present discussion, the medical sectors have used it to monitor disease and assist big data in clinical decision making. Yet the potential harms of big data technology should not be overlooked, especially when individuals' personal lives are being affected significantly. When big data is used to define and construct identity, as in defining who a healthy citizen or employee is, issues of privacy and personal data protection, discrimination and exclusion, and procedural fairness are inevitably involved (Gutwirth and Hildebrandt 2010). One common fear related to identification

from health and medical data is insurance discrimination based on disease susceptibility (Mittelstadt and Floridi 2016, 316). Another fear is group-level of harm from analysis of aggregated data.[64] This is considered to be more problematic as the stigmatization of the affected groups affects all members of the community, not only those who have given consent for their data to be used. Anonymized data subjects maybe grouped according to geographical, socio-economic, ethnic or other characteristic (Mittelstadt and Floridi 2016, 318). Indigenous groups have raised concerns on the risk that they will be singled out for discrimination in big data health research (Greely 2001).

Scholars have advocated for the incorporation of risk and harm assessment to tackle the problems of re-identification and discrimination in data privacy protection. Although they are writing in the larger context of cloud computing and big data technology, their proposed models on data driven accountability is equally applicable in our context of big data health research.[65]

A. Risk Assessment of the Disclosure and Reuse of Data

To ensure accountability, regulation of disclosure and reuse of personal data, including de-identified data,[66] is necessary because third parties may identify the individuals concerned through data combination. This may lead to profiling, and the risks and adverse effects of profiling through data mining and data combination are well recognised.[67] Data brokers have been collecting, analysing, selling and linking

[64] (Mittelstadt and Floridi 2016) (Greely 2001)
[65] Part of the below discussion has been published in (Cheung 2015)
[66] De-identified data includes anonymized and pseudonymous data. The former refers to the personal data that been collected, altered or processed in such a way that it can no longer be attributed to the data subject; while the later refers to explicit identifier being replaced with codes. (Cheung 2015, 78–79)
[67] (Federal Trade Commission (US) 2016) (European Data Protection Supervisor 2015)

individual identities without our knowledge for some time.[68] For example, Acxiom, largest data broker in the US and a marketing giant, holds an average of 1500 pieces of information on each of more than 200 million Americans.[69] Also, it is estimated that each piece of information that users post on Facebook is worth five US cents and that each Facebook user is worth US$100 as a source of information.[70] At present, there is limited regulation of the secondary use of data in most jurisdictions, particularly when they take the ostensible form of de-identified non-personal data. Ultimately, this is an issue of data security, relating to the obligations of data controllers to protect against unauthorised data access, use and disclosure by third parties.

I am not advocating for a complete ban on the use of de-identified data. Indeed, there are legitimate reasons to reuse de-identified (pseudonymous) data,[71] such as in pharmaceutical trials and medical data research or for other legitimate purposes that serve the public interest. In such cases, scholars have recommended that clear guidelines be set, with minimum standards established for the de-identification of datasets and independent reviews of the risk of re-identification before data disclosure.[72] Many have advocated that a specific model be used to measure the continuum of risk involved. For example, Hon et al. and others use the "realistic risk of identification" as a benchmark, [73] whereas Schwartz and Solove suggest the

[68] The FTC uses the term 'data broker' to refer to those that 'collect and aggregate consumers' personal information from a wide range of sources and resell it for an array of purposes, such as marketing, verifying an individual's identity, and preventing financial fraud'. (Federal Trade Commission (US) 2013).
[69] (Kroft 2014).
[70] (Mayer-Schönberger and Cukier 2013, 119)
[71] Article 4(5) defines pseudonymisation to be "the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."
[72] (Hon, Millard, and Walden 2011, 215).
[73] (Hon, Millard, and Walden 2011, 224).

"substantial risk of being identified"(Schwartz and Solove 2011, 1882). More concretely, Ohm recommends that any risk assessment take account of (1) the data-handling techniques used by database owners; (2) the nature of information release, with the public disclosure of data being subject to stricter scrutiny; (3) the quantity of data involved; (4) the likely motives and economic incentives for anyone to re-identify the data; and (5) the trust culture in a particular industry or sector, that is, the existing standard of fiduciary duty or duty of confidentiality in that sector (Ohm 2010). Furthermore, as data identification and combination technologies are advancing at a rapid pace, I contend that any risk assessment concerned should be carried out on a regular basis with citizen scientists being involved, rather than only at the data collection and de-identification and disclosure stages.

B. Re-identification: Data Quality and Size

When considering threats of re-identification external to the original research team or organizations, data quality and size also need to be taken into account. Data quality refers to the nature, sensitivity and linkability of data to individuals (Ohm 2010, 1766), with the latter referring to the different degrees of data identifiability or levels of effort required to identify an individual. An example of good-quality data is the information presented in Google Flu Trends. Regardless of whether its predictions are accurate (Arthur 2014), the information that Google gathers from the online web search queries submitted by millions of individuals are abstracted at a high level and safely aggregated.[74]

Another important element of data quality is data size. The size of a database is determinative of how easy it is to link the information therein to an individual. The

[74] (Schwartz and Solove 2011, 1882)

larger in size it is, the easier that link is to make. However, the law seems to be silent on how much data data controllers may collect, how long they may retain them, what data may or may not be combined and whether stricter security measures are needed for large databases. Ohm argues that new quantitative limits and guidelines should be enacted to address these issues (Ohm 2010, 1767). Such limits and guidelines would undoubtedly have an impact on bio-banks given the vast quantity of data stored in there, but are certainly deserving of further consideration.

In sum, in determining the likelihood of re-identification, we must also consider the quality and quantity of the data in question. Ultimately, the core issue in personal data protection is identity protection.

C. Sensitive Data and Re-combination of Data

Many of the foregoing measures are dependent on the compliance framework of the data controllers and the organisations or companies concerned because participants often have no idea that their data are being re-used and processed or that they have been re-identified. It is therefore important to formulate an alternative privacy framework that is based less on consent and more on holding data controllers accountable for a particular reuse of data based on risk and the likely adverse impact (harm) on data subjects when the unauthorised disclosure and use of data takes place. Namely, the European Union has a higher standard on the use of sensitive personal data which includes genetic, biometric and health data under the new GDPR, while the US regulates the combination of data.

The EU affords sensitive personal data special protection. Recital 51 of the GDPR specifies that personal data "which by their nature, particularly sensitive in relation to fundamental rights and freedom merit specific protection as the context of

their processing could create significant risks to the fundamental freedoms or privacy" should not be processed unless the data subject gives his explicit consent. Although the categories of sensitive data are likely to be controversial in different contexts and cultures,[75] the "sensitive" nature of certain data reveals the underlying values and harm concerned. For example, data related to an individual's health (particularly sensitive health information such as HIV status) may lead to discrimination against that individual.[76] Bearing in mind the threat of harm arising from the re-identification of certain data, organisations need to ensure that sensitive data, which may perhaps be better described as critical data, are stored separately from the general network. They also need to ensure that access to such data is carefully monitored, that their combination with other data cannot easily take place and that their public disclosure is impossible.

Rather than imposing a high standard on a discrete category of sensitive data, there is special restriction on the combination of data in the US. For instance, in 2013, California amended its law on personal information to include regulation of the practice of data combination by imposing new requirements on the operators of commercial websites or online services that collect the personal information of Californian consumers.[77] Under amended s.1798.29 of the California Civil Code, the definition of personal information has been expanded to include an individual's first name or first initial and last name in combination with any one or more of five stated

[75] Article 9 of the EU GDPR prohibits the processing of personal data revealing "racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of genetic data, biometric data for the purposes of uniquely identifying the natural person, data concerning health or data concerning a natural person's sex life or sexual orientation," unless data subjects provide their explicit consent or other conditions under article 9(2) are satisfied.

[76] It was recently reported that the US Federal Bureau of Investigation (FBI) has been collecting racial and ethnic information to map communities around the country based on crude stereotypes about which groups commit various types of crime. (Gangadharan and Vitka 2014).

[77] Personal Information: Privacy SB-46 (2013-2014) reg sess, ch 396, Cal Stat (enacted) <http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201320140SB46> accessed 28 March 2014.

categories of data fields if any is unencrypted: (1) social security number, (2) driver's licence number or California identification card number, (3) bank account number or credit or debit card number in combination with any required security code, access code or password that would permit access to an individual's financial records, (4) medical information and (5) health insurance information.[78] The definition also now includes a username or email address in combination with a password or security question and answer that would permit access to an online account.[79] All of this information is subject to a specific duty of notice of breach and security requirement.[80] California's approach to regulating the combination of certain categories of unencrypted information constitutes a move in the right direction.

The emphasis is rightly on non-attribution to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organisational measures to ensure non-attribution.[81] In addition, the US Federal Trade Commission (FTC) has recommended a more robust system of de-identification and accountability.[82] Rather than toiling with various concepts of de-identified data (anonymous, anonymised and pseudonymous data), the FTC acknowledges that the de-identification of data is not fool proof, and thus there is always a possibility that individuals will be re-identified. Accordingly, it recommends that companies robustly de-identify, and publicly commit to making no attempts to re-identify, data, and contractually require the same public commitment from any downstream users with which they share information. Such requirements should extend to the sharing of data with third parties owing to the possibility of subsequent attribution by later parties.

[78] Cal Civ Code s 1798.29(g)(1).
[79] Cal Civ Code s 1798.29(g)(2).
[80] Cal Civ Code ss 1798.29 and 1798.82.
[81] Recital 29 of the GDPR governs the measures on the use of pseudonymous data.
[82] (Federal Trade Commission (US) 2012, 21–22).

Conclusion

Big data and information communication technologies hold great promises for health and medical citizen science. Citizen scientists can connect and exchange data with one another and with researchers. This has led to growing expectations to access and re-use the data in bio-banks and repositories. In grappling with the shifting nature of data and ever-evolving technology, various notions of consent has been formulated to resolve the tension between researchers' need for data and subject's will for privacy and self-determination. Yet all the attempts to re-fine and re-define consent have proved to be futile conquest of individual's full autonomy.

Embedded in big data analytics is the use of data and personal data, and the matching of data sets. Arguably one does not have enough data and medical science literacy to give meaningful consent to research involving such technology. To most participants, their consent may have reproduced their dependency on expert researchers, medical or state authorities. Regardless whether it is open or dynamic and with variations in between, consent gives only an illusion of control in the big data age. Despite the fact that the formulations of consent may be compatible with existing legal standards, they may be far-cry from ethical imperatives to personal dignity, equality and democratic accountability. Its layered meanings often come with a broader shift of unsolicited responsibility from public healthcare authorities, commercial actors or institutional researchers to individual participants. Regulators have warned that the transfer of risk and regulation should not "signal a reduction in the standards and quality of healthcare provision"(European Group on Ethics in Science and New Technologies to the European Commission 2015, 62). What lies

behind seemingly empowerment of citizens should not be a disguised form of exploitation of extraction and sale of personal data, which may lead to discrimination against individuals or groups.

While consent is still essential in medical and health research, it must be assured by a complementary system of data driven accountability. Consent alone is not enough to restore autonomy to individual and citizen scientists in fast evolving data-intensive research. As there are different dimensions and forms of citizen science, so should the participation of citizen scientists at various stages of research following the life cycle of data usage including risk and harm assessment, re-identification are-combination of data. The solution to attain autonomy must come through a comprehensive set of citizen science practice involving data research.

Bibliography

Angrist, Misha. 2009. “Eyes Wide Open: The Personal Genome Project, Citizen Science and Veracity in Informed Consent.” *Personalized Medicine* 6 (6): 691–99. doi:10.2217/pme.09.48.

Arthur, Charles. 2014. “Google Flu Trends Is No Longer Good at Predicting Flu, Scientists Find.” *The Guardian*, March 27, sec. Technology. https://www.theguardian.com/technology/2014/mar/27/google-flu-trends-predicting-flu.

Article 29 Data Protection Working Party. 2011. “Opinion 15/2011 on the Definition of Consent.” http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf.

Auffray, Charles, Rudi Balling, Inês Barroso, László Bencze, Mikael Benson, Jay Bergeron, Enrique Bernal-Delgado, et al. 2016. “Making Sense of Big Data in Health Research: Towards an EU Action Plan.” *Genome Medicine* 8 (June). doi:10.1186/s13073-016-0323-y.

Ball, Madeleine P., Jason R. Bobe, Michael F. Chou, Tom Clegg, Preston W. Estep, Jeantine E. Lunshof, Ward Vandewege, Alexander Wait Zaranek, and George M. Church. 2014. “Harvard Personal Genome Project: Lessons from Participatory Public Research.” *Genome Medicine* 6: 10. doi:10.1186/gm527.

Beauchamp, Tom L., and James F. Childress. 2012. *Principles of Biomedical Ethics*. UK: Oxford University Press. https://global.oup.com/ushe/product/principles-of-biomedical-ethics-9780199924585?cc=hk&lang=en&.

Bhutta, Zulfiqar A. 2004. “Beyond Informed Consent.” *Bulletin of the World Health Organization* 82 (10): 771–77.

Caulfield, Timothy, Ross EG Upshur, and Abdallah Daar. 2003. “DNA Databanks and Consent: A Suggested Policy Option Involving an Authorization Model.” *BMC Medical Ethics* 4: 1. doi:10.1186/1472-6939-4-1.

Cheung, Anne SY. 2015. “Re-Personalizing Personal Data in the Cloud.” In *: Privacy and Legal Issues in Cloud Computing*, 69. UK: Edward Elgar. https://www.elgaronline.com/view/9781783477067.00011.xml.

Council of Europe. 1997. *Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine* *. https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168007cf98.

European Data Protection Supervisor. 2015. “Opinion 7/2015 Meeting the Challenges of Big Data.” https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2015/15-11-19_Big_Data_EN.pdf.

European Group on Ethics in Science and New Technologies to the European Commission. 2015. “The Ethical Implications of New Health Technologies and Citizen Participation.” Opinion No. 29. Brussels. http://ec.europa.eu/research/ege/pdf/opinion-29_ege.pdf.

Evans, Barbara J. 2016. “Barbarians at the Gate: Consumer-Driven Health Data Commons and the Transformation of Citizen Science.” *American Journal of Law and Medicine* 4 (42). https://papers.ssrn.com/abstract=2750347.

Federal Trade Commission (US). 2012. “Protecting Consumer Privacy in an Era of Rapid Change: Recommendations For Businesses and Policymakers.” USA. https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-

commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

———. 2013. “What Information Do Data Brokers Have On Consumers, And How Do They Use It.” Prepared Statement Before the COMMITTEE ON COMMERCE, SCIENCE, & TRANSPORTATION UNITED STATES SENATE. https://www.ftc.gov/sites/default/files/documents/public_statements/prepared-statement-federal-trade-commission-entitled-what-information-do-data-brokers-have-consumers/131218databrokerstestimony.pdf.

———. 2016. “Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues (FTC Report).” January. https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf.

Felt, Ulrike, Milena D. Bister, Michael Strassnig, and Ursula Wagner. 2009. “Refusing the Information Paradigm: Informed Consent, Medical Research, and Patient Participation.” *Health (London, England: 1997)* 13 (1): 87–106. doi:10.1177/1363459308097362.

Gangadharan, Seeta Peña, and Sean Vitka. 2014. “Knowing Is Half the Battle: Combating Big Data’s Dark Side Through Data Literacy.” *Slate*, April 2. http://www.slate.com/blogs/future_tense/2014/04/02/white_house_big_data_and_privacy_review_we_need_federal_policy_about_digital.html.

Greely, Henry T. 2001. “Informed Consent and Other Ethical Issues in Human Population Genetics.” *Annual Review of Genetics* 35 (1): 785–800. doi:10.1146/annurev.genet.35.102401.091453.

Gutwirth, Serge, and Mireille Hildebrandt. 2010. “Some Caveats on Profiling.” In *Data Protection in a Profiled World*, edited by Serge Gutwirth, Yves Poullet, and Paul De Hert, 31–41. Springer Netherlands. http://link.springer.com/chapter/10.1007/978-90-481-8865-9_2.

Hon, W. Kuan, Christopher Millard, and Ian Walden. 2011. “The Problem of ‘Personal Data’ in Cloud Computing: What Information Is Regulated?—the Cloud of Unknowing.” *International Data Privacy Law* 1 (4): 211–28. doi:10.1093/idpl/ipr018.

Irwin, Alan. 1995. *Citizen Science: A Study of People, Expertise and Sustainable Development*. USA: Routledge.

Kaye, Jane, Edgar A Whitley, David Lund, Michael Morrison, Harriet Teare, and Karen Melham. 2015. “Dynamic Consent: A Patient Interface for Twenty-First Century Research Networks.” *European Journal of Human Genetics* 23 (2): 141–46. doi:10.1038/ejhg.2014.71.

Kroft, Steve. 2014. “The Data Brokers: Selling Your Personal Information.” *CBS News*, March 9. http://www.cbsnews.com/news/data-brokers-selling-personal-information-60-minutes/.

Kuehn, Bridget M. 2013. “Groups Experiment With Digital Tools for Patient Consent.” *JAMA*. August 21. http://jamanetwork.com/journals/jama/fullarticle/1730508.

Mathews, Debra J.H., and Leila Jamal. 2014. “Revisiting Respect for Persons in Genomic Research.” *Genes* 5 (1): 1–12. doi:10.3390/genes5010001.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt.

Meier, Benjamin. 2002. “International Protection of Persons Undergoing Medical Experimentation: Protecting the Right of Informed Consent.” *Berkeley Journal of International Law* 20 (3): 513. doi:doi:10.15779/Z38693R.

Mittelstadt, Brent Daniel, and Luciano Floridi. 2016. “The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts.” *Science and Engineering Ethics* 22 (2): 303–41. doi:10.1007/s11948-015-9652-2.

Mont, Marco Casassa, Vaibhav Sharma, and Siani Pearson. 2012. “EnCoRe: Dynamic Consent, Policy Enforcement and Accountable Information Sharing within.” http://www.hpl.hp.com/techreports/2012/HPL-2012-36.pdf.

Ohm, Paul. 2010. “Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization.” *UCLA Law Review*, no. 57: 1701.

O’Neil, Onora. 2002. *Autonomy and Trust in Bioethics*. UK: Cambridge University Press.

Ploug, Thomas, and Søren Holm. 2016. “Meta Consent – A Flexible Solution to the Problem of Secondary Use of Health Data.” *Bioethics* 30 (9): 721–32. doi:10.1111/bioe.12286.

Prainsack, Barbara Prainsack. 2014. “Understanding Participation: The ‘Citizen Science’ of Genetics.” In *Theory, Technology and Society: Genetics as Social Practice: Transdisciplinary Views on Science and Culture*, edited by Barbara Prainsack Prainsack and Gabriele Silke Werner-Felmayer, 147. Routledge.

Rothstein, Mark A., John T. Wilbanks, and Kyle B. Brothers. 2015. “Citizen Science on Your Smartphone: An ELSI Research Agenda.” *Privacy Laws and International Biobank Research*, no. Winter: 897.

Schmietow, Bettina. 2016. “Ethical Dimensions of Dynamic Consent in Data-Intense Biomedical Research—Paradigm Shift, or Red Herring?” In *Ethics and Governance of Biomedical Research, Research Ethics Forum*, edited by D Strech and M Mertz. Springer. https://www.researchgate.net/publication/302595936_Ethical_Dimensions_of_Dynamic_Consent_in_Data-Intense_Biomedical_Research-Paradigm_Shift_or_Red_Herring.

Schwartz, Paul, and Daniel Solove. 2011. “The PII Problem: Privacy and a New Concept of Personally Identifiable Information.” *New York University Law Review*, no. 86(January): 1814.

Seife, Charles. 2016. “23andMe Is Terrifying, but Not for the Reasons the FDA Thinks.” *Scientific American*. Accessed November 22. https://www.scientificamerican.com/article/23andme-is-terrifying-but-not-for-the-reasons-the-fda-thinks/.

Shirk, Jennifer L., Heidi L. Ballard, Candie C. Wilderman, Tina Phillips, Andrea Wiggins, Rebecca Jordan, Ellen McCallie, et al. 2012. “Public Participation in Scientific Research: A Framework for Deliberate Design.” *Ecology and Society* 17 (2). doi:10.5751/ES-04705-170229.

Sloot, Bart van der, and Sascha van Schendel. 2016. “Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study.” *JIPITEC* 7 (2). http://www.jipitec.eu/issues/jipitec-7-2-2016/4438.

Sreenivasan, Gopal. 2003. “Does Informed Consent to Research Require Comprehension?” *Lancet* 362 (December): 2016–18.

Steinsbekk, Kristin Solum, Bjørn Kåre Myskja, and Berge Solberg. 2013. “Broad Consent versus Dynamic Consent in Biobank Research: Is Passive Participation an Ethical Problem?” *European Journal of Human Genetics* 21 (9): 897–902. doi:10.1038/ejhg.2012.282.

Sweeney, Lataya. 1997. “Weaving Technology and Policy Together to Maintain Confidentiality.” *The Journal of Law, Medicine & Ethics* 25 (2–3): 98–110, 82.

Tallacchini, Mariachiara, Philip Boucher, and Susana Nascimento. 2014. “European Commission, JRC Science and Policy Reports, Emerging ICT for Citizens’ Veillance.” file:///C:/Users/anne/Downloads/Emerging%20ICT%20for%20Citizens%20Veillance.pdf.

Teare, Harriet, M. Morrison, Edgar A. Whitley, and Jane Kaye. 2015. “Towards ‘Engagement 2.0’: Insights from a Study of Dynamic Consent with Biobank Participants.” *Digital Health*. http://dhj.sagepub.com/.

“Tech Companies See Market Opportunity in Healthcare Innovation.” 2016. *Financial Times*. Accessed November 22. https://www.ft.com/content/709aa784-efd4-11e4-ab73-00144feab7de.

The White House. 2015. “FACT SHEET: President Obama’s Precision Medicine Initiative.” *Whitehouse.gov*. January 30. https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative.

Wilbanks, John T. 2014. “Portable Approaches to Informed Consent and Open Data.” In *Privacy, Big Data and the Public Good*, 234. USA: Cambridge University Press.

Williams, Hawys, Karen Spencer, Caroline Sanders, David Lund, Edgar A Whitley, Jane Kaye, and William G Dixon. 2015. “Dynamic Consent: A Possible Solution to Improve Patient Confidence and Trust in How Electronic Patient Records Are Used in Medical Research.” *JMIR Medical Informatics* 3 (1). doi:10.2196/medinform.3525.