

ARTICLE

# Who should have access to genomic data and how should they be held accountable? Perspectives of Data Access Committee members and experts

Mahsa Shabani<sup>\*1</sup>, Adrian Thorogood<sup>2</sup> and Pascal Borry<sup>1</sup>

Facilitating the responsible access to genomic research data is an emerging ethical and scientific imperative. Data Access Committees (DACs) assess the ethical footing and scientific feasibility of the data access requests and evaluate the qualification of applicants to ensure they are *bona fide* researchers. Through semi-structured interviews, we explored the opinions and experiences of 20 DAC members and experts concerning the users' qualification criteria and mechanisms to hold users accountable. According to our respondents, such evaluation is necessary to ensure applicants are trustworthy, meet a certain level of expertise or experience and are aware of the rules and the associated concerns with genomic data sharing. The respondents noted, however, that the qualification criteria are fragmented or are poorly delineated at times. Thus, developing qualification criteria seems vital for an objective, fair and responsible access procedure. Similarly, the access review will benefit from using common ways of verifying the users' affiliations. Furthermore, some DAC members expressed concern over the uncertain oversight of downstream data use, in particular where data are shared across borders. DAC members and experts did not consider current sanctions and enforcement procedures to be crystal clear. Therefore, data sharing policies should address this gap by establishing proportionate sanctions both against data producers and data users' non-compliance. Users' home institutes will need to have an active role in keeping oversight on the downstream data uses, considering their ultimate responsibility if wrongdoings happen.

*European Journal of Human Genetics* (2016) 24, 1671–1675; doi:10.1038/ejhg.2016.111; published online 24 August 2016

## INTRODUCTION

Sharing genomic research data is an emerging ethical and scientific imperative. Funding bodies, research organizations and journals increasingly require researchers to share the results of genomic studies with a wider range of users in an easily accessible manner. Data sharing is considered crucial, to optimize the use of data sets generated by the support of public money and strengthen the statistical power of data. International and national policies and guidelines have provided overarching frameworks to steer data-sharing practices such as the Genomic Data Sharing Policy<sup>1</sup> issued by the National Institutes of Health (NIH) in the United States and the report on Governance of Data Access by the Expert Advisory Group on Data Access in the United Kingdom.<sup>2</sup> The MalariaGEN, the Wellcome Trust Case Control Consortium and the International Cancer Genome Consortium are examples of research consortia that have embarked on data sharing.<sup>3,4</sup>

Despite the perceived benefits, sharing individual-level genomic data has triggered a number of concerns for research participants.<sup>5</sup> Notably, privacy of the data subjects could be endangered, as the insufficiency of the traditional approaches such as de-identification of genomic data sets has been demonstrated.<sup>6–8</sup> Individuals have voiced concerns over potential harmful uses of data that could result from privacy breaches such as discriminatory uses by employers or insurance companies.<sup>9</sup> Despite scarce evidence of privacy breaches so far, a cautionary approach towards data sharing has been favored, to minimize the likelihood of incidents and to maintain public trust in

research institutions.<sup>10,11</sup> In addition, the downstream uses of data bring the discussion regarding adequate consent mechanisms into focus. In particular, a one-off consent may not sufficiently encompass all aspects of data sharing. This has led to some discussions on alternative approaches to consent, to enable the ongoing involvement of the participants in the process of research.<sup>12,13</sup>

Furthermore, concerns have been raised regarding adequate acknowledgement of data producers.<sup>14–16</sup> Researchers are worried about receiving credit for sharing their hard-won data and also expect mutual benefit to accrue from their data-sharing efforts.<sup>17</sup> Such concerns have been reiterated in a number of policy statements and reports such as the Fort Lauderdale Agreement 2003, which states 'The contributions and interests of the large-scale data producers should be recognized and respected by the users of the data.'<sup>18</sup> In response, some mechanisms such as setting a publication moratorium to respect publication priorities of the data producers have been adopted. Nevertheless, the sufficiency and effectiveness of those policies are subject to discussion.<sup>19,20</sup>

A number of different ways are emerging through which users' access to these databases may be managed, including open and controlled access mechanisms. In the latter format, access requests to data sets are assessed for the purpose of approval or disapproval by the corresponding Data Access Committee members (DACs) to ensure responsible downstream uses.<sup>21</sup> The mandate of DACs is typically to assess the ethical and scientific merits of access requests, and the qualifications of applicants. A review of the current structure

<sup>1</sup>Center for Biomedical Ethics and Law, Department of Public Health and Primary Care, University of Leuven, Leuven, Belgium; <sup>2</sup>Center of Genomics and Policy, Department of Genetics, McGill University, Montreal, QC, Canada

\*Correspondence: M Shabani, Center for Biomedical Ethics and Law, Department of Public Health and Primary Care, University of Leuven, Kapucijnenvoer 35 block d, Box 7001, B-3000 Leuven, Belgium. Tel: +32 1637 9517, Fax: +32 1633 6952; E-mail: mahsa.shabani@med.kuleuven.be

Received 1 March 2016; revised 7 July 2016; accepted 19 July 2016; published online 24 August 2016

of the existing DACs yet reveals a significant variety. Although some of the DACs are established in an institution or consortium level with formal meetings and established operating procedures, others are located in the small research groups, mainly composed of principle investigator(s).<sup>3,4,22,23</sup> The DACs associated with the database of Genotypes and Phenotypes (dbGaPs) in the United States and the European Genome-phenome archive (EGA) exemplify the existing practices.<sup>24,25</sup>

Surprisingly, the adequacy of access review by the DACs in line with the ultimate goals of controlled-access model has received little scrutiny to date. To bridge this gap, through a qualitative study we investigated the experiences and attitudes of the DAC members and experts on the components of access review, the existing tools and mechanisms to facilitate access review and the adequacy of such access review in fulfilling the goals of controlled-access model of data sharing. In a previous article, we reported the perspectives of the DAC members and experts on reviewing the ethical aspects and scientific merits of the proposed uses and adequacy of the pertinent tools and mechanisms.<sup>26</sup> The results showed the interviewees were ambivalent about the scope and rigor of such review by DACs and the adequacy of available tools and mechanisms such as consent forms, data access agreements and guidelines to achieve the goals of review. Here we report the attitudes and experiences of the respondents with regards to assessment of applicant qualifications and the effectiveness of existing mechanisms to monitor and respond to violations of data-use conditions.

## MATERIALS AND METHODS

We conducted 20 semi-structured interviews with key informants. Sixteen semi-structured interviews were conducted with members of DACs involved in reviewing access requests for genomic data available in databases of EGA and dbGaP. Purposive sampling was used. We consulted the lists of DACs from the EGA and the dbGaP to retrieve a DAC's contact information, which directed us to a DAC member. In few instances, the contacted DAC member referred us to another DAC member(s) with more experience. In addition, we interviewed four experts in the field who had either published in this field or were members of advisory committees related to data sharing (Table 1). Interviewing experts assisted us to gather insights from people who were not involved in a committee, but given their advisory role or expertise in the field had opportunity to reflect on the potential shortcomings and the pertinent solutions. Invitation letters were sent by e-mail and interviews were conducted via telephone, Skype or in person between November 2014 and May 2015, and were audio-recorded. Audio files were anonymized and transcribed verbatim. Preliminary coding was conducted by MS and discussed within a team for validation. The final analysis of the transcripts was performed by MS using NVivo 10 software by QSR International and inductive content analysis methodologies, followed by further discussions and development of main themes together with PB and AT. The Social and Societal Ethics Committee of University of Leuven approved this study in October 2014.

## RESULTS

DAC members presented their experiences reviewing applicants' qualifications and discussed existing access conditions. Respondents identified perceived shortcomings of current procedures and suggested ways for improvements. They also highlighted the importance of maintaining oversight of data use and discussed the potential role of DACs and other bodies such as home institutions in this process.

### Assessment of the qualification of the applicants

*Experiences.* The assessment of the applicants' qualifications is highlighted as a core component of access review in most interviews. The main purpose of assessing applicants' qualification is to ensure they are actually 'bona fide' researchers, aware of the rules of responsible data use.

*'We are basically there to make sure that people are bonafide scientific researchers, that they are real scientific researchers, that they want to do scientific research and that they want to do it, respecting the basic rules of art.'* (Consortium DAC member, Interview 5)

Nevertheless, respondents noted the concept of a *bona fide* researcher and who is qualified as a *bona fide* researcher are surrounded with uncertainty. Checking the affiliation of the applicants constitutes a significant part of assessment of the data access requests. According to the respondents, this will let the DACs to verify whether the applicants are associated with a credible institute that vouches for them. In addition, when there is any limitation on the type of researcher accessing data (eg, no commercial purposes), checking the affiliation of the applicants will be necessary.

*'It is mainly about being associated with a credible academic institution. Because, the Data access agreement has to be signed not only by the scientists, but also by the institution.'* (Consortium DAC member, Interview 9)

Limited mechanisms are available to systematically check the affiliation of applicants. Checking the e-mail address of applicants or 'Googling' their academic profile were mentioned as examples of current practices.

*'So it is the case that I have always got requests from people or almost always got requests from people with e-mails that are linked to universities or research institution...and so when that's the case, then to me it is pretty obvious that it is a research purpose. So I haven't examined it in more detail.'* (Single Research Group DAC member, Interview 3)

**Table 1 Overview of the interviewees by location, professional/educational background and type of DAC**

Location/number of interviews	Professional/educational background	Types of DACs
Europe (10)	Bioinformatics (7)	Single research group (5)
North America (9)	Biology, Biochemistry, Genetics (9)	Consortium (8)
Australia (1)	Ethics (2)	Institution (3)
	Law (2)	Experts
		Have publication records in the field (2)
		Member of advisory committees on data sharing (2)
Total: 20		

Occasionally, DAC requested extra information from applicants, such as the record of publications, as a demonstration of expertise in the field or seniority. The expertise of the users is believed to indicate their awareness of the pertinent data use concerns. Therefore, some demonstration of knowledge and expertise seemed necessary.

*'Because in some ways it is a case that you want the end user to understand the issues. [In order] to be really confident that they understand why some of the issues might be in place or why some of the requirements are in place. In which case then some demonstration of knowledge and expertise does seem necessary. But how high that bar is, I don't think it has to be exceptionally a high bar.'* (Single Research Group DAC member, interview 3)

**Challenges.** DAC members highlighted various requirements in terms of affiliation of the applicants. For instance, some DACs required applicants to possess an academic affiliation or a permanent academic position, but not others.

*'So one thing that we decided is if you are coming from an academic institution, you need to be a full faculty member, meaning a professor, and to be able to be independently requesting your fund from funding organization. So meaning student can't get access to the data. They need to go through their professor, so the professor can apply for the student. That would work. I am involved with some other organizations and they are like, no we want the students to have access to the data, so not everybody agrees on those things.'* (Consortium DAC member, interview 5)

Some DAC members were concerned about such criteria, as it may pose a barrier to access by non-academics or PhD students. According to one respondent, such requirements may become contrary to the original goals of data sharing, namely providing an easy and broad access to the data sets.

*'What happens about researchers who don't have a current affiliation to an institute? What happens to the students who have great ideas? The great thing about human genome [Project] has been that anybody who has a smart idea in the middle of the night could download the data and could do some analysis.'* (Institution DAC member, Interview 1)

**Potential solutions.** The importance of adopting common procedures was pronounced by some of our respondents. Such common access procedures are argued to be beneficial, to avoid redundancies and to enable qualified applicants to access a variety of data sets without seeking authorization for each database separately.

*'It would be handy if several institutions and organizations had a common way of vetting the users. If there was a standardized way of doing so. [Meaning] If you are already proving that to one repository, then you shouldn't need to do it again. Because then you could say: I am already certified by these guys and you don't need to certify me again and it saves lot of time and you don't need to check it anyway.'* (Expert, Interview 16)

### Oversight on downstream uses

**Experiences.** The interviewees were asked about the robustness of the ongoing monitoring of data use after access was granted. DAC members and experts did identify a few existing oversight

mechanisms, such as requiring users to provide interim or final reports. Reporting the final publications out of analysis of the data and any Intellectual Property claims have been required in some occasions.

*'We cannot really police and go and check to see on daily basis that investigators have done something with the data that they should not have. So a lot of time you hear about these incidents either through word of mouth or at the time of reporting period.'* (Institution DAC member, Interview 11)

**Challenges.** DAC members and experts noted that oversight mechanisms are generally limited. In addition, granting access often to multiple international applicants was mentioned as an underlying reason that makes oversight hard if not impossible.

*'You are allowing use in many other countries around the world, there is no point in thinking you can actually go and catch individuals somewhere else and hold them responsible for what they have done.'* (Consortium DAC member, Interview 2)

Respondents pointed out, however, that the proportionality of the requirements remains an important consideration. For instance, some DAC members questioned the significance of such reporting, claiming this is redundant, or burdensome for the users, particularly when they use multiple data sets.

*'That is more burdensome to actual users of data, because users of data need to submit annual reports on data use and when a lab has multiple projects, using multiple datasets, it is annoying and completely redundant and unnecessary work [to report].'* (Consortium DAC member, Interview 7)

**Potential solutions.** As a result of sharing data via databases such as dbGaP and EGA users could transfer data to their system and potentially use data for various purposes. This brings into question the feasibility and robustness of oversight, as our respondents noted. In response, alternative models of data sharing were suggested by some respondents, which let users' access to data in a protected environment without downloading data.

*'What could be done is instead of moving the data you move the computation to the data and then they can record what you do, like you really record the transaction. And then maybe they can also get better sense of what exactly you do with all of the data. It will not be perfect but just that...if you know someone is watching, this already takes 90 percent of malicious uses away.'* (Expert, Interview 20)

Presumably, such a model could allow the current access review process to be simplified. As one DAC member discussed, this allows monitoring of the use without approval in advance and avoiding overly conservative access decisions. Arguably, this could lessen the burden on the DACs and allow an effective monitoring.

*'So as I said I don't think we entirely know because there hasn't been enough experience and I think we are going to evolve towards a model where the touch of authorization is lighter because you can do monitoring in parallel. So essentially I think until now the decisions of DACs and their considerations has been somehow heavier because of this feeling that if you approve something you lose control...that's kind of psychological aspect of it and if you then move to an environment where you haven't lost control as you can always stop people accessing, then*

*people can feel slightly more relaxed about allowing people to explore data in different ways.’ (Institution DAC member, Interview 1)*

### Sanctions

*Experiences and challenges.* The respondents underlined a need for sanctions against breach of data use conditions in downstream uses of data. Indeed, in the absence of meaningful sanctions, the effectiveness of the access review procedure was questioned. For this purpose, the responsibility of home institutes was underscored. This aligns with the current approach towards the data access agreements, where co-signing the agreement by the home institutes is obligatory. Given the users’ distance from the data producers, the home institutes seem to be in a better standing to monitor data use and to respond to potential violations.

*‘You have to have some organization that you can hold responsible and the best organization is the employer, because that employer risks if they allow a misuse of data. If that happens it is the employer who can be black-listed, who can be fined, who can be brought into disrepute, who can be held responsible for the conduct. So there is a distinction between ensuring the person has an appropriate qualification, but also in my view, more importantly, ensuring that there is someone who is responsible for the conduct of the researcher.’ (Consortium DAC member, Interview 2)*

*Potential solutions.* Respondents discussed professional codes of conduct as an instrument to set up the sanctions, although not all researchers are governed by a professional code. In addition, the sanctions could be grounded in general regulatory penalties.

*‘I think that will be most effective if it comes through professional codes of conduct, but I think that if we think outside of data sharing to how genomic data are used within society, it might be helpful to think about what general civil penalties might be for misuse of the data. As I think the potential probably is greater outside of research community than it is within the research community.’ (Institution DAC member, Interview 12)*

### DISCUSSION

The underlying reasons for evaluating qualification of the applicants and the associated challenges with the current practices have been discussed by our respondents. The DACs evaluate the qualification of the users, to ensure they are *bona fide* researchers. According to the respondents of this study, such evaluation is necessary, to ensure users are trustworthy, meet a certain level of expertise or experience and are aware of the rules and the associated concerns with genomic data sharing. The recent report on Governance of Data Access by the Expert Advisory Group on Data Access reiterates the significance of such evaluation: ‘Those who donate their data for research probably expect that the research will be carried out by those with demonstrated competence in the field. The research is also more likely to be useful and productive if carried out by someone with adequate research experience, and there are accepted (although not always enforceable) standards of good conduct amongst professional researchers.’<sup>2</sup> Our respondents noted, however, that the qualification criteria are fragmented or are poorly delineated at times. Although holding a permanent academic position seemed to be an access requirement for some DACs, others believed this created an unjustified barrier. Similarly, some DACs but not others required users to provide the record of publication as an indication of seniority and also awareness of the rules.

Developing qualification criteria thus seems vital for an objective, fair and responsible access procedure.<sup>27</sup> As the DAC members highlighted though, such criteria should be proportionate to the associated risks.

In addition, the affiliation(s) of applicants should be verified. The purpose of this verification is twofold. The first purpose of verifying affiliations is to ensure the home institute could be held responsible in case of wrongdoings by the user and in turn will hold the user accountable. In order to achieve the first objective, co-signing data access agreement by the home institute would be necessary.<sup>28</sup> The second purpose is to verify the applicant is affiliated with a trusted institute. Naturally, shifting the assessment of trustworthiness from the researcher to the institution raises the question of what criteria must an institution meet? As with individual users, it may be more difficult to assess the trustworthiness of an institution when it is outside of known collaborations, or in a foreign jurisdiction. Likewise, using the same criteria for non-academic institutes could be a subject of discussion. In turn, the development of clear and standard criteria for trusted institutions is desirable. Such criteria could include applying adequate security and data protection measures, educating and training the employees concerning the best practices, implementing comprehensive data-sharing policies and monitoring the compliance. Moreover, the procedure to verify the affiliation of the applicants is not straightforward at the moment. For instance, consulting available online information about the users via general search is reported as a common practice among DACs for that matter. This underscores a need for adopting robust ways in verifying the users’ affiliations, to streamline the procedure and improve its reliability. Similarly, some respondents suggested a common certification (following universal criteria) for researchers for access to a variety of data sets would be beneficial. To this end, data access rules ‘must be compatible with each other’ and ‘must comply with laws and regulations of relevant jurisdictions’ as Contreras and Reichman put it.<sup>29</sup> Arguably, this will mitigate the burden of several reviews of the qualification of applicants by DACs and overcome the latency in access to the data sets. Unique researcher IDs<sup>30</sup> could potentially serve to provide some degree of stability and reassurance, although the robustness of such a mechanism for the purpose of qualification assessment is questionable.

Maintaining oversight on data users and downstream uses of data has also appeared as a concern for some interviewees. Cross-border and/or cross-institution sharing of data are believed to hamper an effective oversight on the data use. This brings a discussion on the best oversight mechanisms and the potential role of DACs into focus. Some DACs require periodical or final reports from users, to monitor the downstream uses of data and to identify the potential violations of access agreements. However, some respondents were skeptical of the necessity of such requirement, as it allegedly adds another layer of bureaucracy.

DAC members discussed a potential role of other stakeholders such as home institutes in maintaining oversight on data users. Home institutes are included, because they are able to impose administrative sanctions for non-compliance effectively compared with the challenge of enforcing statutory penalties or access contracts, especially across borders.

Nevertheless, proportionate sanctions should supplement the oversight mechanisms. Our respondents did not consider current sanctions and the enforcement procedures to be crystal clear. Similarly, ‘very few formal or informal sanctions for data sharing noncompliance’ have been reported to exist if ‘a scientist fails to share as required or expected.’<sup>31</sup> Data-sharing policies therefore should address this gap by establishing proportionate sanctions both against data producers and data users’ non-compliance. Guidelines for establishing compliance policies are provided by the ‘Accountability Policy’ of the Global

Alliance for Genomics and Health, a consortium of organizations from research, healthcare and industry committed to promoting responsible genomic data sharing.<sup>32</sup> This Policy builds on the Global Alliance's Framework for Responsible Sharing of Genomic and Health Related Data.<sup>33</sup> As a first step to enhance accountability, and in turn the trustworthiness of data-sharing initiatives, the Global Alliance recommends that compliance policies be developed by data stewards and host institutions that specify exactly what constitutes inappropriate behavior (eg, data misuse), what range of responses may be taken and what criteria will be used to assess the severity of a given response. Recognizing that data-sharing standards are emerging, a range of facilitative and punitive responses may be appropriate. Facilitative responses for data stewards could include warnings and additional training. Punitive responses (sanctions) range from reporting to other stakeholders (eg, home institutes, ethics boards, regulatory bodies and other data stewards), suspending or terminating access (for data stewards), or suspending or terminating employment (home institutes). The compliance policy should be made widely available and information on non-compliance should be collected and shared. As accountability mechanisms depend on clear rules and predictable consequences, standardization of such mechanisms across the international research community is desirable.

In addition, some of our respondents suggested oversight could be facilitated by new data-sharing models that allow analysis but retain the data in a secure computational environment.<sup>34</sup> Other approaches allow a limited amount of information to be shared in response to queries made by the users,<sup>35</sup> although their use may be yet limited to specific models of data access. Given the alternative models are not yet widely established, the current model of data sharing seems here to stay for the foreseeable future, as some of our respondents noted.

In conclusion, our study showed the DAC members and experts are ambivalent about the effectiveness and consistency of the current access review procedures, and oversight process. Further investigations are necessary, to identify oversight mechanisms commensurate with the associated risks and concerns. Regardless, structured collaboration between various involved bodies such as DACs and home institutions is crucial in establishing robust oversight mechanisms and responding to data misuses. Moreover, harmonization of access review procedures could be achieved by developing international policies and guidelines that delineate qualification criteria and processes for verifying users' affiliations.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank all DAC members and experts who participated in this study and shared their invaluable experiences with us. This work is kindly supported by the Interfaculty Council for Development Co-operation (IRO) of the University of Leuven and the Research Fund Flanders.

- 4 Parker M, Bull S, de Vries J, Agbenyega T, Doumbo OK, Kwiatkowski DP: Ethical data release in genome-wide association studies in developing countries. *PLoS Med* 2009; **6**: e1000143.
- 5 Shabani M, Bezuidenhout L, Borry P: Attitudes of research participants and the general public towards genomic data sharing: a systematic literature review. *Expert Rev Mol Diagn* 2014; **14**: 1053–1065.
- 6 Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y: Identifying personal genomes by surname inference. *Science* 2013; **339**: 321–324.
- 7 Homer N, Szeleringer S, Redman M *et al*: Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet* 2008; **4**: e1000167.
- 8 Rothstein MA: Is deidentification sufficient to protect health privacy in research? *Am J of Bioeth* 2010; **10**: 3–11.
- 9 Haga SB, O'Daniel J: Public perspectives regarding data-sharing practices in genomics research. *Public Health Genomics* 2011; **14**: 319–324.
- 10 Lemke AA, Wolf WA, Hebert-Beirne J, Smith ME: Public and biobank participant attitudes toward genetic research participation and data sharing. *Public Health Genomics* 2010; **13**: 368–377.
- 11 The Nuffield Council on Bioethics. *A Review of Evidence Relating to Harm Resulting from Uses of Health and Biomedical Data* 2014. Available at: <http://nuffieldbioethics.org/wp-content/uploads/FINAL-Report-on-Harms-Arising-from-Use-of-Health-and-Bio-medical-Data-30-JUNE-2014.pdf>.
- 12 Kaye J, Whitley EA, Lund D *et al*: Dynamic consent: a patient interface for twenty-first century research networks. *Eur J Hum Genet* 2014; **23**: 141–146.
- 13 Erlich Y, Williams JB, Glazer D *et al*: Redefining genomic privacy: trust and empowerment. *PLoS Biol* 2014; **12**: e1001983.
- 14 Kaye J, Heeney C, Hawkins N, de Vries J, Boddington P: Data sharing in genomics-research scientific practice. *Nat Rev Genet* 2009; **10**: 331–335.
- 15 Fecher B, Friesike S, Hebing M: What drives academic data sharing? *PLoS One* 2015; **10**: e0118053.
- 16 Wallis JC, Rolando E, Borgman C L: If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PLoS One* 2013; **8**: e67332.
- 17 van Panhuis WG, Paul P, Emerson C *et al*: A systematic review of barriers to data sharing in public health. *BMC Public Health* 2014; **14**: 1144.
- 18 Report of a meeting organized by the Wellcome Trust and held on 14–15 January 2003 at Fort Lauderdale, USA. *Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility*. 2003. Available at: <https://www.genome.gov/pages/research/wellcomereport0303.pdf>.
- 19 Joly Y, Dove ES, Kennedy KL *et al*: Open science and community norms Data retention and publication moratoria policies in genomics projects. *Med Law Int* 2012; **12**: 92–120.
- 20 Contreras JL: NIH's genomic data sharing policy: timing and tradeoffs. *Trends Genet* 2015; **31**: 55–57.
- 21 Shabani M, Knoppers BM, Borry P: From the principles of genomic data sharing to the practices of data access committees. *EMBO Mol Med* 2015; **7**: 507–509.
- 22 Dyke SO, Hubbard TJ: Developing and implementing an institute-wide data sharing policy. *Genome Med* 2011; **3**: 60.
- 23 Manolio TA, Rodriguez LL, Brooks L *et al*: New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 2007; **39**: 1045–1051.
- 24 Paltou DN, Rodrigues LL, Feolo M *et al*: Data use under the NIH GWAS Data Sharing Policy and future directions. *Nat Genet* 2014; **46**: 934–938.
- 25 Lappalainen I, Almeida-king J, Kumanduri V *et al*: The European Genome-phenome Archive of human data consented for biomedical research. *Nat Genet* 2015; **47**: 692–695.
- 26 Shabani M, Borry P: "You want the right amount of oversight": interviews with data access committee members and experts on genomic data access. *Gen Med* 2016; e-pub ahead of print 21 January 2016; doi:10.1038/gim.2015.189.
- 27 Shabani M, Dyke SO, Joly Y, Borry P: Controlled access under review: improving the governance of genomic data access. *PLoS Biol* 2015; **13**: e1002339.
- 28 Joly Y, Zepps N, Knoppers BM: Genomic databases access agreements: legal validity and possible sanctions. *Hum Genet* 2011; **130**: 441–449.
- 29 Contreras JL, Reichman JH: Sharing by design: data and decentralized commons. *Science* 2015; **350**: 1312–1314.
- 30 ORCID. 2014. Available at: <http://orcid.org/>.
- 31 Pham-Kanter G, Zinner DE, Campbell EG: Codifying collegiality: recent developments in data sharing policy in the life sciences. *PLoS One* 2014; **9**: e108451.
- 32 Regulatory and Ethics Working Group Global Alliance for Genomics and Health. *Accountability Policy*. 2016. Available at: [https://genomicsandhealth.org/files/public/Accountability%20Policy\\_FINAL%20v1\\_Feb10.pdf](https://genomicsandhealth.org/files/public/Accountability%20Policy_FINAL%20v1_Feb10.pdf).
- 33 Knoppers BM: Framework for responsible sharing of genomic and health-related data. *HUGO* 2014; **8**: 3.
- 34 Wallace SE, Gaye A, Shoush O, Burton PR: Protecting personal data in epidemiological research: DataSHIELD and UK Law. *Public Health Genomics* 2014; **17**: 149–157.
- 35 Philippakis AA, Azzariti DR, Beltran S *et al*: The matchmaker exchange: a platform for rare disease gene discovery. *Hum Mut* 2015; **36**: 915–921.

1 NIH. Genomic Data Sharing Policy 2014. Available at: [https://gds.nih.gov/PDF/NIH\\_GDS\\_Policy.pdf](https://gds.nih.gov/PDF/NIH_GDS_Policy.pdf).

2 Expert Advisory Group on Data Access. *Governance of Data Access* 2015. Available at: <https://wellcome.ac.uk/sites/default/files/governance-of-data-access-eagda-jun15.pdf>.

3 Joly Y, Dove ES, Knoppers BM, Bobrow M, Chalmers D: Data sharing in the post-genomic world: the experience of the International Cancer Consortium (ICGC) Data Access Compliance Office (DACO). *PLoS Comput Biol* 2012; **8**: e1002549.