

# Variant data sharing by clinical laboratories through public databases: consent, privacy and further contact for research policies

Mahsa Shabani, LL.M, PhD<sup>1</sup>, Stephanie O. M. Dyke, PhD<sup>2</sup>, Luca Marelli, PhD<sup>3</sup> and Pascal Borry, PhD<sup>1</sup>

Access to detailed variant data is key to inform and verify the interpretation of genomic data. Clinical laboratories can play a significant role in sharing patients' data through public variant databases. To facilitate data sharing, various public databases, such as ClinVar and DECIPHER have been established, which accept data submission from laboratories, clinicians, researchers, and patients. Despite clear benefits to sharing, questions may arise about the adequate form of consent to be obtained from patients when sharing data from their clinical tests through public databases. In this paper, we provide an overview and critical

analysis of the relevant consent policies of the major public databases, and of the consent forms of clinical laboratories that share variant data via ClinVar.

*Genetics in Medicine* (2018) <https://doi.org/10.1038/s41436-018-0316-x>

**Key words:** variants; clinical laboratories; data sharing; public databases; consent

## INTRODUCTION

Access to detailed variant data is key to inform and verify the interpretation of genomic data. To facilitate access to variant interpretation data, the role of clinical laboratories in sharing patients' data through public variant databases has been recognized by professional communities.<sup>1</sup> In a recent statement, the American College of Medical Genetics and Genomics stressed the importance of clinical and laboratory data sharing to improve health care: "Responsible sharing of genomic variant and phenotype data will provide the robust information necessary to improve clinical care and empower device and drug manufacturers that are developing tests and treatments for patients."<sup>2</sup>

Furthermore, increasing support from various professionals and experts towards data sharing by clinical laboratories has been shown in several studies. For instance, results of focus group studies with experts and stakeholders including representatives from regional clinical genetics laboratories across the United Kingdom indicated that, "the majority felt that professionals, clinicians and laboratory scientists would be failing in their duty of care to patients by not sharing data outside individual hospital trusts."<sup>3</sup> Similarly, a survey study with genetic counselors in the United States has concluded:

"Genetic counselors have the ability to support laboratories who publically share their variant data by choosing to order their patients' testing from those laboratories. The majority of clinical genetic counselors responding (79.01%) reported being aware of whether or not the laboratories they order testing from participate in data sharing."<sup>4</sup>

To facilitate data sharing, various public databases that accept data submission from laboratories, clinicians, researchers, and patients, have been established. According to a recent policy draft issued by the Food and Drug Administration, public variant databases would: "(1) operate in a manner that provides sufficient information and assurances regarding the quality of source data and its evidence review and variant assertions; (2) provide transparency regarding its data sources and its operations, particularly around how variant evidence is evaluated and interpreted; (3) collect, store, and report data and conclusions in compliance with all applicable requirements regarding protected health information, patient privacy, research subject protections, and data security; and (4) house sequence information generated by validated methods."<sup>5</sup>

ClinVar and DECIPHER (the Database of Genomic Variation and Phenotype in Humans Using Ensembl

<sup>1</sup>Center for Biomedical Ethics and Law, Department of Public Health and Primary Care, University of Leuven, Leuven, Belgium; <sup>2</sup>Montreal Neurological Institute, McGill University, Montreal, Canada; <sup>3</sup>Life Sciences and Society Lab, Centre for Sociological Research, University of Leuven, Leuven, Belgium.

Correspondence: Mahsa Shabani ([Mahsa.shabani@kuleuven.be](mailto:Mahsa.shabani@kuleuven.be))

Submitted 21 June 2018; accepted: 10 September 2018

Published online: 08 October 2018

Resources) are two major public databases that are frequently used by laboratories for data sharing. ClinVar, which was launched in 2013 and is maintained by the US National Center for Biotechnology Information, is a critical resource for the community that “serves as a primary site for deposition and retrieval of variant data and annotations” and accepts submissions of variants and supporting evidence by researchers, clinical laboratories, expert groups, clinicians, and patients. As of March 2018, a total number of 600,610 submissions from 925 data submitters to ClinVar was reported. DECIPHER is a United Kingdom-based database that was initiated in 2004 and is “a web-based platform for secure deposition, analysis, and sharing of plausibly pathogenic genomic variants from well-phenotyped patients suffering from rare genetic disorders.” Currently, DECIPHER contains data from 25,697 patients who have given consent for broad data sharing. In addition to ClinVar and DECIPHER, other databases, such as the Leiden Open Variant Database, along with locus-specific variant databases and disease-centered variant databases, also allow for data submission of variants.<sup>6</sup>

An example of a disease-centered variant database is the BRCA Challenge (<http://brcaexchange.org/>)—an international data-sharing initiative that is pooling data on BRCA1/2 variants and their clinical interpretation from multiple public resources to provide a comprehensive public resource and advance our understanding of the genetic basis of breast cancer, ovarian cancer, and other diseases. It is an exemplar of what is achievable through aggregation of all data sources.

To date, commitment to open data sharing has been announced by both commercial and public clinical laboratories.<sup>7</sup> For instance, commercial companies such as Counsyl, Illumina, Ambry, and Pathway Genomics have issued statements and highlighted the significance of variant sharing for better patient care, particularly for the interpretation of variants of uncertain significance.<sup>8–11</sup> In addition, to further facilitate data sharing, some laboratories have developed online platforms, such as AmbryShare and Clinvitae to publicly share aggregated genomic data from testing with other laboratories, researchers, and clinicians. Similarly, in the United Kingdom, a survey of National Health Service regional clinical laboratories and specialist laboratories reported that 87% (13/15) of laboratories that responded to the survey currently deposited some data to databases of genetic variants.<sup>3</sup>

However, despite clear benefits to sharing, questions may arise about the adequate form of consent to be obtained from patients when sharing data from their clinical tests through public databases. How much information should be communicated to the patients regarding this data sharing? Should patients have the option to opt out? Under what conditions could laboratories share more detailed information (e.g., supporting individual-level data) about the patients than simply the aggregate variant-level classifications? To address these questions, we provide an analysis of the relevant consent policies of the two major public databases already described—

namely ClinVar (with guidance provided by ClinGen) and DECIPHER—and of the consent forms of 17 clinical laboratories that, as of June 2018, meet the minimum requirement of data sharing as determined by ClinGen (<https://www.clinicalgenome.org/lablist/>). We will discuss this under three main issues (that is, consent, privacy, and the risks of re-identification and further contact with patients for clinical or research purposes) and conclude by providing further analysis on the adequacy of the current policies and approaches.

## CONSENT FOR DATA SHARING

In the framework of ClinVar, the submitter is responsible for determining when consent is required (<https://www.ncbi.nlm.nih.gov/clinvar/docs/submit/>). To guide submitters, ClinGen, in partnership with ClinVar, describes points to consider when determining whether consent is required in relation to the amount of data to be shared. They assert that explicit consent for sharing de-identified variant-level information obtained by laboratories during the course of fee-for-service clinical testing is not necessary.<sup>12</sup> However, consent requirements may change when sharing “more specific individual-level information, such as the distinct phenotypes of each individual observed in a particular laboratory’s experience with a variant”. ClinGen recommends that submitting laboratories discuss sharing more specific information with a relevant Institutional Review Board using a template protocol for data sharing prepared by ClinGen.<sup>12</sup> The template protocol includes a request for consent waiver, due to the fact that the risks of re-identifications are very low and the data are not collected in the research setting. In addition, the template protocol states that the information about the laboratory’s intent to submit summary variant-level information to ClinVar will be made available to patients via “statements on the laboratory test requisition form, test results, and laboratory website.”

Requirement for consent may also change due to a model of data sharing through public databases. In a guidance document provided by DECIPHER, explicit consent for data sharing has been required only when data are shared in an open-access fashion. In contrast, when data submitters could share pseudonymized data within a controlled-access network and with the approved users, explicit consent is not required. Notably, such a controlled-access model is often used in the context of sharing data collected in the research setting. DECIPHER also provides sample assent and consent forms (in nine languages) and information packs for families about data sharing to be used by data submitters.<sup>3</sup>

Looking at the consent forms from the studied clinical laboratories, four major approaches to consent regarding data sharing for clinical and research purposes are observed in practice (see Table 1). The first approach includes explicitly informing patients about sharing their data with other clinical laboratories and/or through public variant databases such as ClinVar, and providing an opt-out option. Only ARUP followed this approach and has prepared a Genetic Testing

**Table 1** Current approaches of the studied clinical laboratories to consent, privacy and further contact

<b>Consent</b>	Specific information about data sharing without opt-out option (Center for Mendelian Genomics, University Medical Center Ljubljana, Color Genomics, GeneDx, Integrated Genetics, Laboratory Corporation of America, Partners Laboratory for Molecular Medicine)	Example: "Test results and submitted clinical information may be shared with other clinical laboratories for the purpose of improving our understanding of the relationship between genetic changes and clinical symptoms. Sharing data in this manner may enable us to provide better interpretations of your genetic findings as well as assist other patients with similar results. We will protect your privacy/confidentiality by replacing your name and other direct identifiers, such as date of birth or medical record number, with a code. The key to the code numbers will be stored securely in the testing laboratory. We will share only de-identified information with outside clinical labs." (Partners Laboratory for Molecular Medicine; <a href="http://personalizedmedicine.partners.org/laboratory-for-molecular-medicine/">http://personalizedmedicine.partners.org/laboratory-for-molecular-medicine/</a> )
	Explicit consent for data sharing with opt-out from data sharing (ARUP)	Example: "In cooperation with the National Institutes of Health's effort to improve the understanding of specific genetic variants, ARUP submits HIPAA-compliant, de-identified (cannot be traced back to the patient) genetic test results and health information to public databases. Confidentiality of each sample is maintained. You may choose not to share your test result by calling ARUP Laboratories at (800) 242-2787 ext. 3301. Your de-identified information will not be shared with public databases after the request is made, but a separate request is required for each genetic test." (ARUP; <a href="https://www.aruplab.com/genetics/resources/consent">https://www.aruplab.com/genetics/resources/consent</a> )
	Specific information about data sharing and research, together with an opt-out option (Ambry, Emory, Invitae)	Example: "Emory Genetics Laboratory retains patient samples indefinitely for validation, educational purposes and/or research. For molecular cytogenetic and some molecular genetic tests, submitted clinical information and test results are also included in HIPAA-compliant, de-identified public databases as part of the National Institute of Health's effort to improve diagnostic testing and our understanding of the relationships between genetic changes and clinical symptoms (for information about the molecular cytogenetic database, visit the consortium website at <a href="https://www.iscaconsortium.org/">https://www.iscaconsortium.org/</a> and for information about the molecular genetic database refer to the individual test descriptions). Confidentiality of each sample is maintained. Patients may request to withdraw consent for the storage of their sample and/or use of the data by [...]." (EMORY; <a href="https://genetics.emory.edu/documents/molecular-testing-consent-form.pdf">https://genetics.emory.edu/documents/molecular-testing-consent-form.pdf</a> )
	No direct reference to data sharing (Athena Diagnostics, Center for Pediatric Genomic Medicine, Children's Mercy Hospital and Clinics, Counsyl, GeneKor MSA, Illumina, Phosphorus Diagnostics LLC, Quest Diagnostics, University of Chicago)	Example: "Athena Diagnostics requests your permission to use your specimen, clinical information and data in a de-identified way (without any identifying information) for research, educational studies, commercial purposes and/or publication, if appropriate." (Athena Diagnostics; <a href="https://www.athenadiagnostics.com/ordering/resources/informed-consent">https://www.athenadiagnostics.com/ordering/resources/informed-consent</a> )
<b>Privacy</b>	De-identification, including the potential risks for re-identification (Ambry, GeneDx, Color Genomics)	Example: "Even though only a code is used for the reporting to the databases, there is a risk that (I/my child) could be identified based on the genetic and health information that is shared. GeneDx believes that this is unlikely, though the risk is greater if I have already shared (my/my child's) genetic or health information with public resources, such as genealogy websites." (GeneDx; <a href="https://www.genedx.com/all-forms/">https://www.genedx.com/all-forms/</a> )

Table 1 continued

**Information about further contact for research** (Arup, Ambry, Counsyl, GeneDx, Partners Laboratory for Molecular Medicine, Phosphorus Diagnostics LLC)

Example: “Rarely, the provider who ordered your genetic test may contact you about participating in research. If you are interested in participating and/or receiving information about research studies, your provider can give you the researcher’s contact information. A researcher will not have access to any of your personally identifying information unless you contact them.” (ARUP; <https://www.aruplab.com/genetics/resources/consent>)

and Data Sharing document, which provides information about data sharing to patients and allows them to opt out of the sharing. The second approach is informing the patients about data sharing but not providing an option for opting out. Five laboratories, such as Partners HealthCare and GeneDX, inform patients about sharing de-identified data with other laboratories or through public databases such as ClinVar, but do not provide an option for opting out. Color Genomics provides an option to opt out from the inclusion of data in its research database, but not from the contribution of de-identified information about genetic variants to public databases. Color Genomics’ website further explains that such data sharing “is not an external research study, but rather a database that is used to uncover more links between genetics and disease.”

In the third approach, the plans for data sharing are discussed, together with the aim for using data and samples for validation, educational purposes and/or research. Three laboratories follow this approach. For instance, Emory discusses these two in its “information for health care providers and patients”, and provides an option to opt out. In a slightly different approach, AmbryShare has developed a separate consent form “Sharing Genomic Data for Discovery”, in which it asks for patients’ explicit consent for a research study called “Sharing Genomic Data for Discovery—Global”, which aims to share de-identified data with researchers worldwide. Thereby, Ambry considered the data sharing itself as part of a larger research endeavor and prepared a separate consent for this purpose.

In the fourth approach, in the consent forms of laboratories such as Athena Diagnostics, data sharing is not mentioned directly; only the use of specimens, clinical information and data in a de-identified way for research, educational studies, commercial purposes, and/or publication is mentioned. Eight laboratories followed this approach.

**PRIVACY AND THE RISK OF RE-IDENTIFICATION**

To address privacy concerns, only variant-level data, and not large genomic datasets, are permitted to be shared through public databases such as DECIPHER and ClinVar. ClinVar also allows submitting case data about each individual with the variant as supporting evidence, as long as the individual is not identifiable according to National Institutes of Health guidelines. According to the ClinVar guidance, the risks of re-

identification of de-identified variant-level data are very low, owing to the fact that full genomic datasets are not shared and multiple variants from the same individual are not linked together in the database.

In publicly disclosing revealing information about the data submitters, ClinVar and DECIPHER adopt different approaches. Currently, ClinVar reveals the name of the submitting laboratories, arguing: “As clinical testing laboratories often receive samples for testing from around the world, and are not necessarily limited to only receiving samples from a local population, naming the submitting clinical laboratory is not believed to put individual privacy at substantial risk—one cannot assume with certainty the location of an individual based on the location of the laboratory performing their testing.” In contrast, DECIPHER does not publicly reveal the global location of the submitting laboratories, although the data submitters can be contacted via a request through the coordinators.

In addition, the consent forms of clinical laboratories inform patients that only their de-identified data will be shared. In three cases, the risks of re-identification are communicated to patients in the consent form. The GeneDX consent form is one example, where the patients are informed that although the risk of re-identification is generally low, the risk is greater if they have already shared their genetic or health information with public resources, such as genealogy websites. The rest of the consent forms do not include the risks of re-identification. In addition, the safeguards provided for individuals by national laws against genetic discriminations in the context of employment and health insurance (e.g., the US Genetic Non-Discrimination Act) are mentioned in most of the consent forms.

Some laboratories also assure patients that data will be shared by federal Health Information Privacy and Accountability Act of 1996 (HIPAA)-compliant databases, or explicitly refer to ClinVar as an example of public databases that data will be shared with.

**FURTHER CONTACT WITH THE PATIENTS FOR RESEARCH PURPOSES**

Sharing variant data through public databases may result in a request for contact with the patient or access to further information about the patient by other laboratories, researchers, or clinicians. It should be noted that in the rare-disease

community, individual laboratories and databases come together early on to share data in the hope of finding similar patients with rare disorders. Public variant databases such as DECIPHER therefore also contribute to distributed data-sharing platforms such as the Matchmaker Exchange (MME).<sup>13</sup> The MME allows genomic discovery through exchange of phenotypic and genotypic profiles. Thereby, users of such networks could also search for the existence of specific variants/genes in databases such as DECIPHER and could follow up and request access to detailed information according to the terms and conditions of each participating database.

DECIPHER's "an introduction for families" leaflet includes the possibility of requests for sharing further information to share with other clinicians who are interested in similar clinical cases: "If your clinician is contacted by another clinician who through DECIPHER has identified other individuals with the same/similar genomic variant and the same/similar clinical features, you may be contacted and asked whether you wish to give permission for further details to be exchanged with a view to furthering understanding of this genomic variant."

The consent forms of six laboratories, such as ARUP, inform patients that they may be contacted later on for participation in research, and that their identifying information will be revealed to interested researchers/clinicians only after they explicitly agree. In this consent, patients are informed about the possibility of further contact, and are given an option to decide later whether they would be willing to share more information or collaborate with the researchers. In other examples, such as Counsyl, the patients are given an option in the consent form to "opt out of such research or future contact". The possibility for further contact for research is not explicitly included in the rest of the consent forms.

### ARE THE CURRENT POLICIES TOWARDS CONSENT, POLICY, AND FURTHER CONTACTS ADEQUATE?

Currently, public variant databases strive to facilitate variant data sharing by hosting data submissions through online platforms, and providing guidance on how to adequately address issues related to consent, privacy, and secondary uses of data. Such guidance is of paramount importance for commercial and public laboratories, to assist them in sharing patients' data in accordance with ethical and legal standards and addressing the following major points.

First, although public variant databases provide general guidance regarding consent and privacy, questions remain about the most adequate policies regarding opt-out options, the amount of information to be communicated to the patients regarding data sharing, and the adopted privacy-preserving methods. One can argue that sharing minimum information, including aggregate variant-level classifications, to improve patient care should not require obtaining explicit consent. However, some consent forms discuss data sharing together with research activity, and thereby adopt an opt-out

option. An underlying reason for adopting such an approach is that publicly sharing data could result in secondary use of data by a broad range of users, thus potentially going beyond patient care. Although it is difficult to draw a fine line between research and clinical use when sharing data through public databases, adopting an opt-out or explicit-consent policy can be considered burdensome for clinical laboratories that are taking extra efforts to share data.

Second, de-identification is used as a mechanism to protect the privacy of data subjects. This seems to adequately address privacy concerns, owing to the fact that very limited data are planned to be shared through public variant databases. However, from a regulatory perspective, the question of whether de-identified genetic data may still be considered personal (i.e., identifiable and non-anonymous) and require compliance by data protection requirements, including obtaining consent, represents a contentious issue that is presently addressed in distinct ways in different jurisdictions. For instance, in the United States, the federal medical Privacy Rule, promulgated under HIPAA, singles out 18 distinct identifiers, the removal of which is said to make the resulting information "not individually identifiable".<sup>14</sup> However, insofar as DNA sequences are not comprised within the list of such identifiers, genetic data may be considered as de-identified information in the absence of other identifiers.<sup>15</sup>

In contrast with the US approach, the European Union's newly enforced (May 2018) General Data Protection Regulation (GDPR) only considers irreversibly de-identified data as non-personal data, which are therefore exempt from compliance with the rules for sharing personal data, including obtaining specific consent. In contrast, reversibly de-identified genetic data may still be considered personal data in the European Union, to be protected by data-protection regulations.<sup>16</sup> However, the question of what constitutes irreversibly de-identified data is not addressed by the GDPR in clearly defined terms. The GDPR states that anonymous data (that is, data that are not uniquely related to an identified or identifiable natural person) should not be considered as personal data and, accordingly, do not fall within the scope of the regulation (GDPR recital 26)<sup>17</sup>. In addition, the GDPR stipulates that personal data should be considered anonymous insofar as the data subject cannot be identified "by any means reasonably likely to be used [...] either by the controller or by any other person" (GDPR recital 26; see also Article 29 Working Party,<sup>18</sup> opinion 05/2014). To ascertain whether means are reasonably likely to be used to identify the natural person, the GDPR further states that "account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments" (GDPR recital 26). Therefore, and in contrast with HIPAA's definite list of identifiers, the GDPR does not aim to enumerate identifiers, nor does it aim to provide a precise definition of what constitutes an "identifier". As such, the GDPR leaves the question open as to when, and at which conditions, genetic

information such as gene variant data can be deemed to be irreversibly de-identified, and thus exempt from compliance with the rules for sharing personal data. In addition, consistent with its decentralized, controller-anchored, and accountability-based approach,<sup>19</sup> the GDPR devolves to controllers the responsibility to address such a question in the context of their processing activities, thus raising significant challenges for the genetic data-sharing endeavor that undoubtedly warrant further scholarly scrutiny.

Third, although implicit or opt-out options for consent are adopted for variant-level sharing, the current guidance provided by variant databases stresses the importance of obtaining explicit consent once more detailed information is to be shared publicly (tiered data sharing), as both the risk of re-identification and the probable use of the data beyond the immediate clinical applications are increased. This is particularly useful when the data discovery (only minimum response to queries based on disease name, structured phenotype descriptions, and names of genes of interest to find similar cases) and detailed data access are separated. For instance, according to MME policy: “The second level of matchmaking involves sharing more detailed genotypic and/or phenotypic data that may be unique or sensitive. [...] Use of this level of information carries a possible risk of re-identification and as such requires appropriate patient consent.”<sup>20</sup> Additionally, to mitigate privacy concerns, public variant databases may provide different levels of data access to general users and registered users, enabling the acceptance of terms and conditions related to an appropriate use of the data for more detailed patient information that presents a higher level of risk.<sup>21</sup> In sharing more data with such protections, it becomes possible to limit research and other potential uses according to the consent preferences of individuals using simple tools such as the Consent Codes, which let registered users know of consent-based restrictions on data use (e.g., use of the data is limited to health/medical/biomedical purposes).<sup>22</sup>

In conclusion, data sharing by clinical laboratories would benefit from further guidance provided by relevant professional communities regarding the management of data sharing and access, and related consent considerations, to facilitate responsible data sharing. In particular, clinical laboratories will benefit from further clarification regarding the nature of data sharing, namely to be considered as a research activity, clinical activity or both, and the implications for the adequate consent policy. In addition, the pertinent guidelines should adopt adequate safeguards for the privacy of the individuals that are scaled to the identified risks associated with sharing genomic data. As sharing variant data is crucial in improving diagnoses for patients, setting unnecessarily restrictive requirements should not be favored.

Although in this paper we have focused on the role of clinical laboratories in data sharing, the crucial role of clinicians in data sharing should not be overlooked. This could open discussions regarding the rights and responsibilities of clinicians and laboratories regarding data sharing, including the custodianship of patient data,

management of consent, and scope of data sharing. The relevant guidelines could be beneficial in providing guidance in this regard.

Finally, patients and their families' views regarding data sharing by laboratories and the adequacy of the current consent policies should be taken into consideration. Previous studies have shown that although individuals generally support data sharing, they have reservations regarding how data can be used, and are concerned about privacy risks and potential misuses of data. Perceptions regarding what is considered an adequate consent policy are varied.<sup>23</sup> The use of complementary methods, such as videos together with written consent forms, in improving the comprehension of informed consent materials has been also discussed in the literature. For instance, a recent study by ClinGen has shown that participants' performance on comprehension questions “significantly improved over baseline after reading the consent form [prepared by ClinGen] and continued to improve after watching the video.”<sup>24</sup> In addition, further empirical evidence is needed to investigate the current experiences of laboratories regarding the management of data sharing and identify the associated challenges.<sup>25</sup>

## ACKNOWLEDGMENTS

M.S. is funded by a postdoctoral fellowship from the Research Foundation–Flanders (FWO). L.M. has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 753531.

## DISCLOSURE

The authors declare no conflicts of interest.

## REFERENCES

- Harrison SM, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet Med*. 2017;19:1096.
- ACMG Board of Directors. Laboratory and clinical genomic data sharing is crucial to improving genetic health care: a position statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19:721–722.
- Raza S, et al. Data sharing to support UK clinical genetics and genomics services. Workshop report. PHG Foundation. 2015. <http://www.phgfoundation.org/report/data-sharing-to-support-uk-clinical-genetics-and-genomics-services>. Accessed 18 September 2018.
- Colleen A, Brokamp E. The utility of genomic variant databases in genetic counseling. *Human Genetics Theses and Capstones*. Paper 15. 2016.
- FDA. Use of public human genetic variant databases to support clinical validity for next generation sequencing (NGS)-based in vitro diagnostics. 2016.
- Brookes AJ, Robinson PN. Human genotype–phenotype databases: aims, challenges and opportunities. *Nat Rev Genet*. 2015;16:702–715.
- Jamal SM, et al. Practices and policies of clinical exome sequencing providers: analysis and implications. *Am J Med Genet A*. 2013; 161:935–950.
- Counsyl. Sharing data is part of Counsyl's DNA. 2015. <https://blog.counsyl.com/2015/06/10/sharing-data-is-part-of-counsyls-dna/>. Accessed 18 September 2018.
- Illumina. Illumina contributes to ClinVar database. 2016. <https://www.illumina.com/company/news-center/press-releases/2016/2213381.html>. Accessed 18 September 2018.

10. Pathway Genomics. Pathway Genomics announces commitment to free the data movement at NSGC conference. 2014. <https://www.pathway.com/press-releases/pathway-genomics-announces-commitment-to-free-the-data-movement-at-nsgc-conference/>. Accessed 18 September 2018.
11. Invitae. Is better patient care dependent on sharing genetic data? 2014. <http://blog.invitae.com/is-better-patient-care-dependent-on-sharing-genetic-data/>. Accessed 18 September 2018.
12. Azzariti DR, et al. Points to consider for sharing variant-level information from clinical genetic testing with ClinVar. *Mol Case Stud*. 2018;4:a002345.
13. Philippakis AA, et al. The Matchmaker Exchange: a platform for rare disease gene discovery. *Hum Mutat*. 2015;36:915–921.
14. The Office for Civil Rights. Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. 2010.
15. Kulynych J, Greely HT. Clinical genomics, big data, and electronic medical records: reconciling patient rights with research when privacy and science collide. *J Law Biosci*. 2017;4:94–132.
16. Shabani M, Borry P. Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation. *Eur J Hum Genet*. 2018;26:149–156.
17. EU General Data Protection Regulation. 2016. <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504>. Accessed 18 September 2018.
18. Opinion 05/2014 on Anonymization Techniques. 2014. [http://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index\\_en.htm](http://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index_en.htm). Accessed 18 September 2018.
19. Marelli L, Testa G. Scrutinizing the EU General Data Protection Regulation. *Science*. 2018;360:496–498.
20. Dyke SO, et al. “Matching” consent to purpose: the example of the Matchmaker Exchange. *Hum Mutat*. 2017;38:1281–1285.
21. Dyke SO, et al. Registered access: a ‘triple-A’ approach. *Eur J Hum Genet*. 2016;24:1676–1680.
22. Dyke SO, et al. Consent codes: upholding standard data use conditions. *PLoS Genet*. 2016;12:e1005772.
23. Shabani M, Bezuidenhout L, Borry P. Attitudes of research participants and the general public towards genomic data sharing: a systematic literature review. *Exp Rev Mol Diag*. 2014;14:1053–1065.
24. Riggs ER, et al. Development of a consent resource for genomic data sharing in the clinical setting. *Genet Med*. 2018. <https://doi.org/10.1038/s41436-018-0017-5>. Accessed 18 September 2018.
25. Thorogood A, Cook-Deegan R, Knoppers BM. Public variant databases: liability? *Genet Med*. 2016;19:838–841.