Building the Medical Information Commons: Case Study on BRCA Variant Data-Sharing Practices

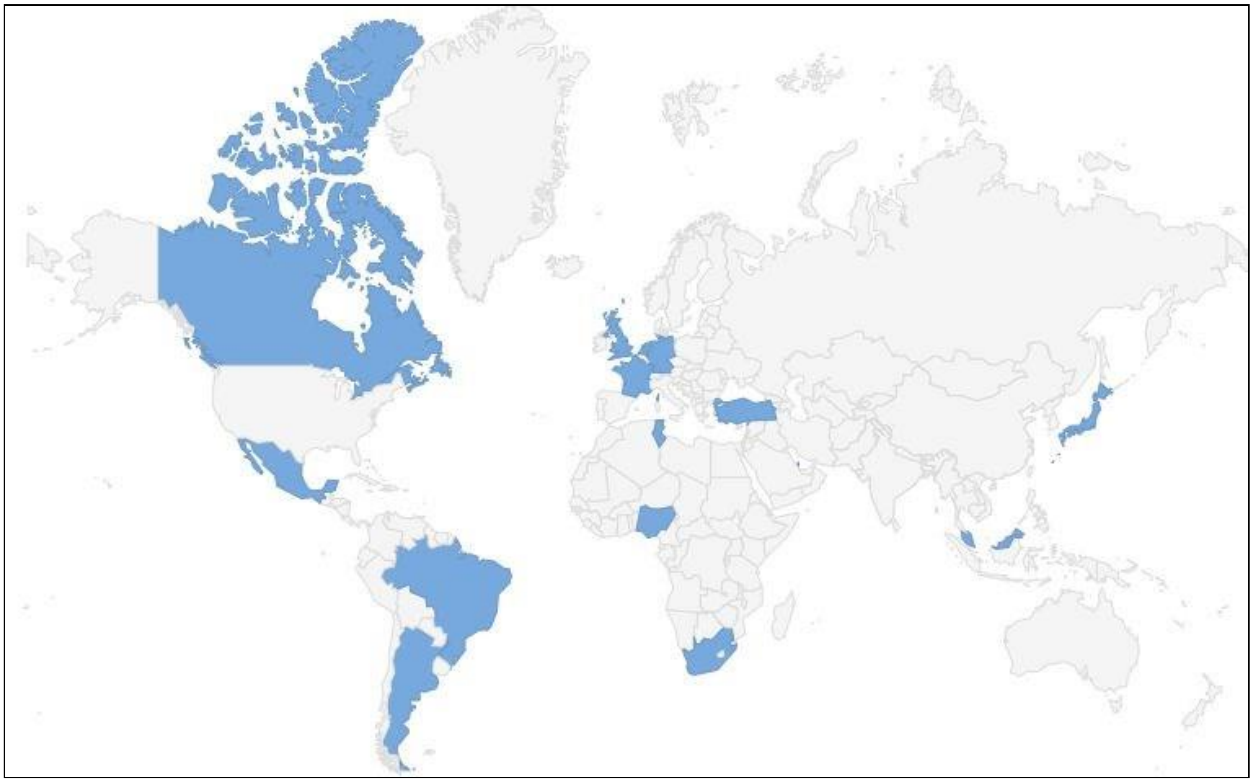
Results from International Database and Laboratory Interviews

Jan 2019

Authors:
Lena Dolman, M.Sc. | Manager, Strategy and Outreach; Global Alliance for Genomics & Health (GA4GH); Toronto, Canada
Rachel G. Liao, Ph.D. | Senior Project Manager; Broad Institute; Cambridge, USA

Table of Contents

Figure 1. Map displaying the 15 interviewed countries in blue.



- Argentina
- Brazil
- Canada
- France
- Germany
- Japan
- Malaysia
- Mexico
- Netherlands
- Nigeria
- Qatar
- South Africa
- Tunisia
- Turkey
- United Kingdom

Executive Summary & Implications

Between November 2017 and February 2018, phone- or email-based interviews were conducted with 10 laboratories and 8 databases from 15 countries across North America, South America, Europe, Africa, the Middle East, and Asia.

The interviews revealed that a wealth of BRCA1/2 variant data is being produced and collected by worldwide institutions of varied scale. Individual labs report providing BRCA testing for anywhere between a dozen to several thousand patients per year, while individual databases include hundreds to thousands of unique BRCA variants. Importantly, a number of centres are amassing data on non-European ancestry individuals, providing insights into populations that have previously been underrepresented in genomics databases.

Such international, diverse datasets are of immense value to the global community of clinicians and scientists tasked with interpreting BRCA variants, and it is therefore important that these datasets are accessible and can be used to derive clinical interpretations. Indeed, substantial support for data sharing was articulated by the interviewed labs: six were already sharing BRCA data, three were preparing to commence sharing soon, and one was open to the idea in future. Among databases, the desire to share is also strong, as all databases were developed in order to support variant sharing to enable improved clinical interpretations. Several databases currently limit sharing to consortium members, while others allow public viewing and download of variants, classifications, and evidence.

However, despite this support for data sharing, key barriers continue to persist. These include a lack of personnel to oversee this task (the most commonly-cited barrier among labs), lack of time, cost constraints, technical constraints, legal constraints and liability concerns, and lack of a sharing culture, among others. At the same time, several factors are motivating sharing: supportive collaborators & peers with which to share (the most commonly-cited incentive among labs), existing tools and databases to facilitate sharing, and the belief that sharing meets patient expectations, improves the quality of work for all, supports more robust clinical decision-making, and helps the community. Interviewed centres mentioned that sharing might be further motivated by the existence of resources (financial and human) dedicated to the sharing task, by education/training on existing tools/databases, and by policy changes (e.g. making data sharing a requirement for lab licensing).

The interviews also provided insights into how to improve clinical databases in order to enhance their utility and further facilitate sharing. Most labs (8 out of 10) routinely access ClinVar during variant interpretation, while several labs also utilize population databases (ExAC, 1000 Genomes) and LOVD. However, due to some of the barriers listed above (e.g. technical, time, and personnel constraints), only two of the 10 labs were contributing data into ClinVar (one directly, and one via the Canadian Open Genetics Repository). In general, the labs articulated preferences for comprehensive, curated, expert-reviewed databases that were easy to use, and that provided transparency on data submitters, data quality, and variant classification methodology.

Despite the consensus on the value of data sharing, the interviews also highlighted areas of disconnect across the global community. Groups are continuing to interpret variants in differing and inconsistent ways, despite the existence of guidelines from expert bodies like ENIGMA and ACMG. Many groups seem to be unaware of existing resources, tools, and community initiatives for sharing, and might therefore benefit from targeted education and awareness efforts. As of yet, no one group has taken on this task of convening international BRCA data producers and users to share experiences and best practices, and to highlight opportunities for data access and sharing. Indeed, the need for training opportunities to increase awareness of existing databases for sharing, and the perceived absence of international collaboration and of a sharing culture, were raised in the interviews.

Overall, our report found that robust BRCA1/2 data sharing is already taking place in both testing laboratories and databases around the world, often despite substantive barriers. There is a general consensus for the value of sharing to support both variant classification and well-informed clinical decision-making among providers and variant carriers. With the right mixture of legal and regulatory support, technological development, and financial reimbursement, alongside social support in the form of education and the convening of worldwide groups, it is our belief that the world's data could be accessed to an even great extent than now, for the further improvement of medical understanding and support of BRCA1/2 variant carriers and their families.

Summary of Key Findings

Between November 2017 and February 2018, we conducted phone- or email-based interviews with 10 laboratories and 8 databases from 15 countries across North America, South America, Europe, Africa, the Middle East, and Asia.

Labs

This section includes the summary results from among 10 international labs interviewed (in Argentina, Brazil, Canada, Malaysia, Mexico, Nigeria, Qatar, South Africa, Tunisia, and Turkey).

Experience with BRCA Testing

Among the interviewed labs (Table 1), six of the ten were currently sharing BRCA data (Argentina, Brazil, Canada, Malaysia, Nigeria, and Tunisia), three planned to commence sharing in the future (Mexico, Qatar, and South Africa), and one lab did not have plans to share BRCA data, but was open to the idea in future (Turkey). Those labs that were currently sharing BRCA data had been sharing for the past 1-20 years.

Table 1. Summary of BRCA data sharing practices across 10 interviewed labs.

Country	Currently Sharing BRCA Data?	Year Started Sharing BRCA Data
Argentina	Yes	2017
Brazil	Yes	After 2015
Canada	Yes	2014
Malaysia	Yes	After 2006
Mexico	No (but plan to in future)	n/a
Nigeria	Yes	1998
Qatar	No (but plan to in future)	n/a
South Africa	No (but plan to in future)	n/a
Tunisia	Yes	2015
Turkey	No (but open to the idea)	n/a

Resources used in BRCA Variant Interpretation

The 10 interviewed labs were asked to specify which resources they routinely access during the process of BRCA variant interpretation (Figure 2 & Table 2).

The most commonly-mentioned resource was ClinVar (routinely used by 8/10 labs, and considered a “go-to” resource by 5 of the labs). Five labs mentioned reviewing literature, 5 used in-house databases, and 4 used population databases (e.g. ExAC, 1000 Genomes). Commercial or paid-access databases - including ThermoFisher’s BRCA Oncomine database, HGMD, and

the Sophia Genetics database - were used by four of the labs. The BRCA Exchange was used by 1 lab (Malaysia).

Figure 2. Resources routinely accessed in BRCA variant interpretation process (among 10 interviewed labs)

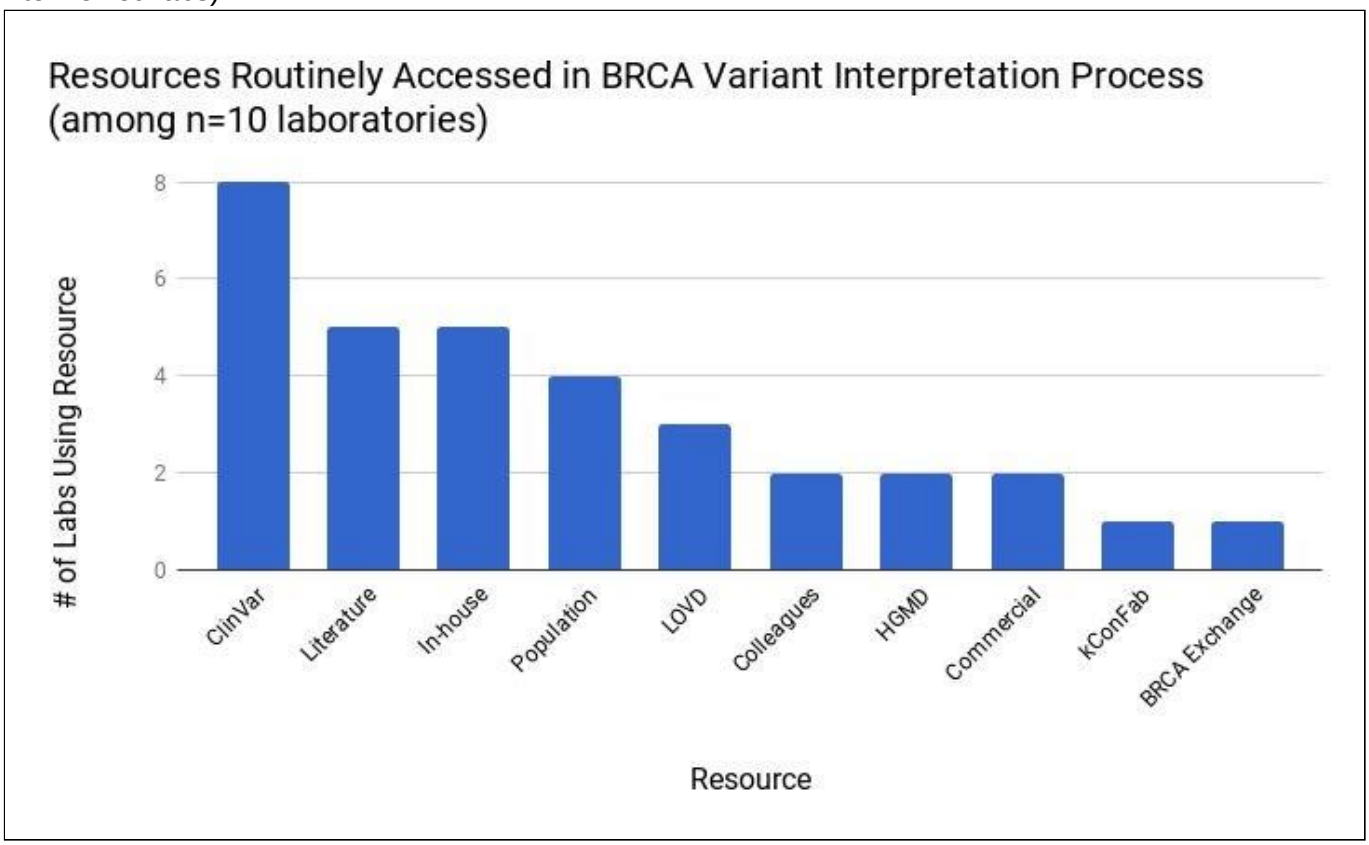


Table 2. Summary, by country, of resources routinely accessed in BRCA variant interpretation process

Country *(* = currently sharing)*	Routinely-Accessed Public Databases	Other Routinely-Accessed Resources
Argentina*	ClinVar, LOVD	Colleagues, literature
Brazil*	ClinVar, population databases	HGMD, internal Latin American database
Canada*	ClinVar, LOVD, kConFab database, IARC, population databases	Colleagues, literature, in-house database
Malaysia*	ClinVar, LOVD, BRCA Exchange	
Mexico	ClinVar, population databases	Literature, in-house database, ThermoFisher BRCA Oncomine database
Nigeria*	ClinVar	In-house database
Qatar	Population databases, disease- and locus-specific databases	Literature, in-house database
South Africa	ClinVar	Sophia Genetics database
Tunisia*	ClinVar	Literature
Turkey	HGMD	

The lab groups also provided positive and negative feedback relating to their experiences with using large, public databases, as well as feedback on databases in general (including commercial databases). This feedback is summarized in Table 3. Common themes that emerged included:

a) A preference for curated databases with review by expert panels;
b) A preference for databases with easy user experience & simple data submission process;
c) A preference for being able to easily determine how variants were classified;
d) A preference for databases that compile information from other sources (e.g. literature references, IARC classifications);
e) A preference for global, ethnically-diverse data sets;
f) An aversion for databases with multiple, differing interpretations for a given variant;
g) An aversion for lack of clarity on: data quality; data submitters; how interpretations were reached.

Table 3. Lab feedback on public databases and other (e.g. commercial) databases used in BRCA variant interpretation

Feedback on Existing Public Databases		
Public Database (# Labs Routinely Accessing)	Positive Feedback	Negative Feedback
ClinVar (8)	• Very comprehensive source for BRCA information • Helpful especially when classifications have been reviewed/contributed by an expert panel • Useful for determining if a variant was reported previously • Simple, easy, efficient, and updated • Clear, user-friendly	• Interpretations from large commercial labs labs 'pollute' data by adding differing interpretations for a given variant • Question data quality, especially from commercial lab submitters • Lack of clarity around pathogenicity classifications • Interpretations are grouped together, rather than separated out by phenotype • Not curated • Time/effort required to prepare data for submission to ClinVar
LOVD (3)	• Includes IARC classifications • Provides references, allowing for easy, independent review of the evidence • Most informative for variant classification	• No central place to access all data • Less information provided for variants • Less comprehensive for BRCA genes (relative to ClinVar) • Hard to determine who deposited data and why
BRCA Exchange (1)	• Excellent, comprehensive, global, and evidence-based resource for BRCA variant[s] • Extensively curated per variant, and updated monthly	
Other Database-Related Feedback		
• Preference for curated databases • When using commercial databases, preference is given to those services that clarify how variant classifications are reached, that make clear reference to the ACMG guidelines, and that allow users to see and compare interpretations made by other customers within the database • An ideal database for community data sharing would be easy to use, curated, and reciprocal; would provide information on other ethnic groups; and would enable easy data sharing as part of lab technicians' daily routine (not requiring substantial additional effort) • Most public data sets predominantly reflect European-ancestry patients; therefore, data is less valuable for other populations		

BRCA Data Sharing: Current Practices, Barriers, and Incentives

Six of the ten labs are currently sharing BRCA data (Argentina, Brazil, Canada, Malaysia, Nigeria, and Tunisia), three plan to commence sharing in the future (Mexico, Qatar, and South Africa), and one lab does not yet have plans to share BRCA data, but is open to the idea (Turkey).

Among the 6 labs currently sharing BRCA data (Table 4), three contribute to CIMBA, two share their data with ClinVar (in the case of the Canadian lab, this sharing is via the Canadian Open Genetics Repository, or COGR), two share with LOVD, and two share with ENIGMA. One group (Malaysia) also shares data with BCAC, ABRCA, and BIC. At least two of the labs share data locally with other research labs or collaborators. The most frequently-shared data elements are variants alongside their interpretations or associated evidence (e.g. publications). The Malaysian lab also shares pedigrees and penetrance data, while the Tunisian lab shares case-level and clinical or epidemiological research data. None of the labs have dedicated personnel involved in data sharing; this task is typically shared among team members and/or lab leadership.

Table 4. Summary of BRCA data sharing practices among 6 sharing laboratories

Country (Year Started BRCA Sharing)	Whom Sharing With	Data Elements Shared	Personnel Involved in Sharing	Consent Includes Language on Data Sharing?
Argentina (2017)	LOVD, ENIGMA	Variants & associated publications	Volunteers, research fellows (no dedicated staff)	Yes (anonymized results can be shared, published)
Brazil (after 2015)	ClinVar	Variants & interpretations	Shared among team (no dedicated staff)	Yes (broad consent, also allowing recontact)
Canada (2014)	COGR (which contributes to ClinVar)	Variants & interpretations	Shared among team (no dedicated staff)	No
Malaysia (after 2006)	CIMBA, ENIGMA, BCAC, ABRCA, BIC, LOVD, collaborators	Variants, pedigrees, penetrance	Shared among team (no dedicated staff)	Yes (allows data use in collaborative research projects while maintaining privacy)
Nigeria (1998)	CIMBA	Variants	Shared among team (no dedicated staff)	Unknown
Tunisia (2015)	Research labs/ collaborators (soon: CIMBA)	Variants, pathogenicity, link to evidence, case-level data, research data (clinical, epidemiological)	Leadership of lab & breast cancer project (no dedicated staff)	Yes (data may be used in research projects, and transferred to other labs in Tunisia or abroad for research purposes)

When asked about barriers to BRCA data sharing, all 10 labs mentioned a lack of personnel as a key barrier (Figure 3 & Figure 4). Other barriers mentioned by at least half of the labs included a lack of time for data sharing activities, and inadequate budget / cost constraints. Five barriers were only mentioned by the subset of labs currently sharing data: technical constraints, liability concerns, lack of education/awareness on available databases for sharing, competitive incentives to keep data, and lack of a sharing culture (Figure 4). An additional four barriers were only mentioned by the labs not currently sharing data: institutional barriers (e.g. needing permission from multiple levels of their institution), logistical barriers, concerns over maintaining data sovereignty (e.g. to avoid potential misappropriation or exploitation of local patient data by other world regions), and potential inability to gain patient consent (e.g. due to lack of available counselling services, low regional patient literacy rates, and patient fears regarding insurance discrimination). Finally, IP / data ownership issues, and legal barriers (against international sharing specifically), were also mentioned by more than one lab.

The lab groups had less to report when asked about incentives or motivations for BRCA data sharing (Figure 5). The most common incentive for sharing was the existence of supportive collaborators/peers/community with which to share. Other incentives mentioned more than once included the existence of tools to initiate data sharing (making the process easier), and a sense that sharing improves the quality of work for all involved (e.g. stronger variant interpretations, and emergence of community-supported best practices). Finally, sharing was seen as meeting patient expectations, supporting patient autonomy (e.g. in the case of returning results directly to patients), and helping the community.

The breakdown of responses per country is given in Table 5, while further descriptions of the barriers & incentives are given in Table 6.

Figure 3. Barriers to BRCA data sharing reported among 10 international laboratories.

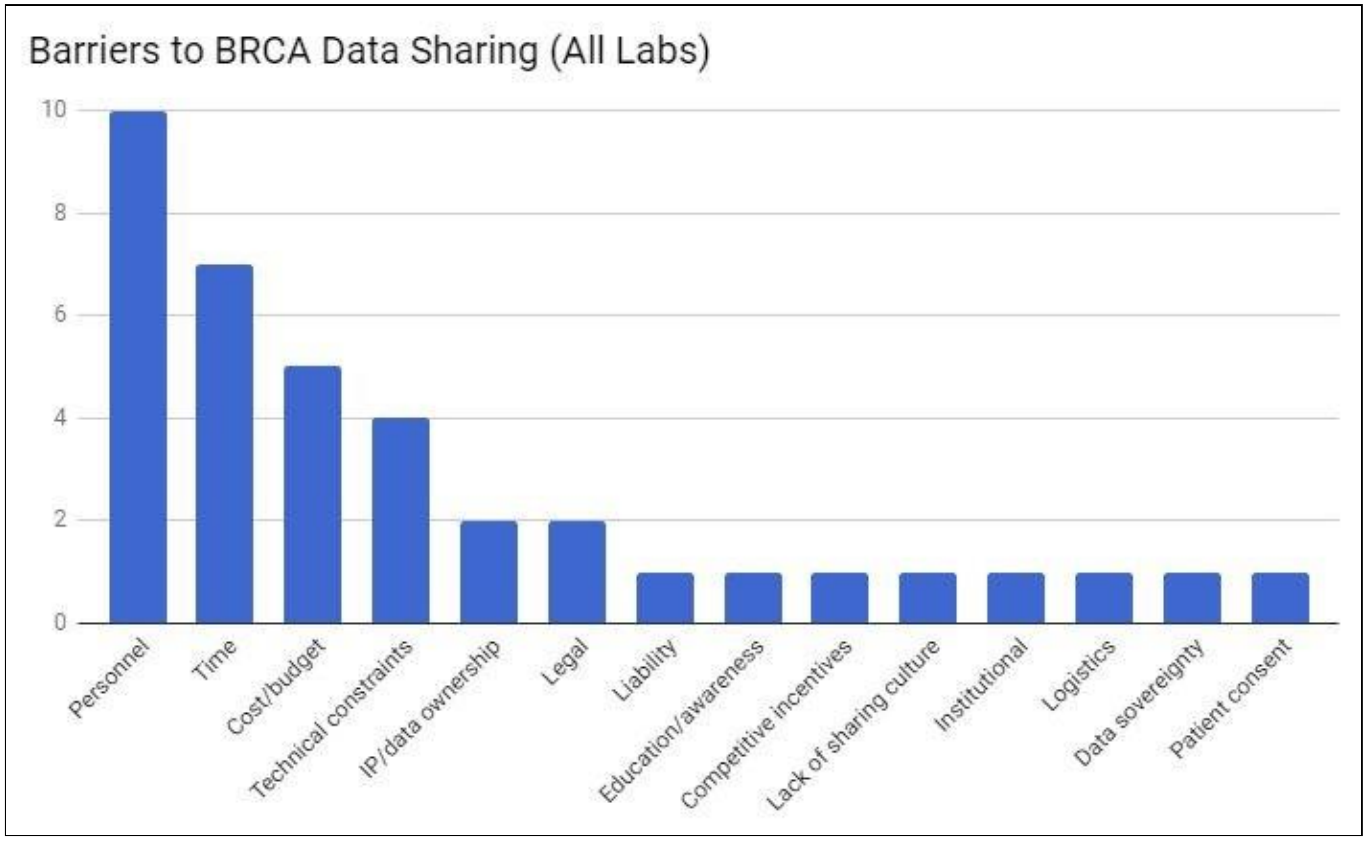


Figure 4. Barriers to BRCA data sharing, segregated by current data sharing practice (i.e. among 6 labs currently sharing data, and 4 labs not currently sharing data)

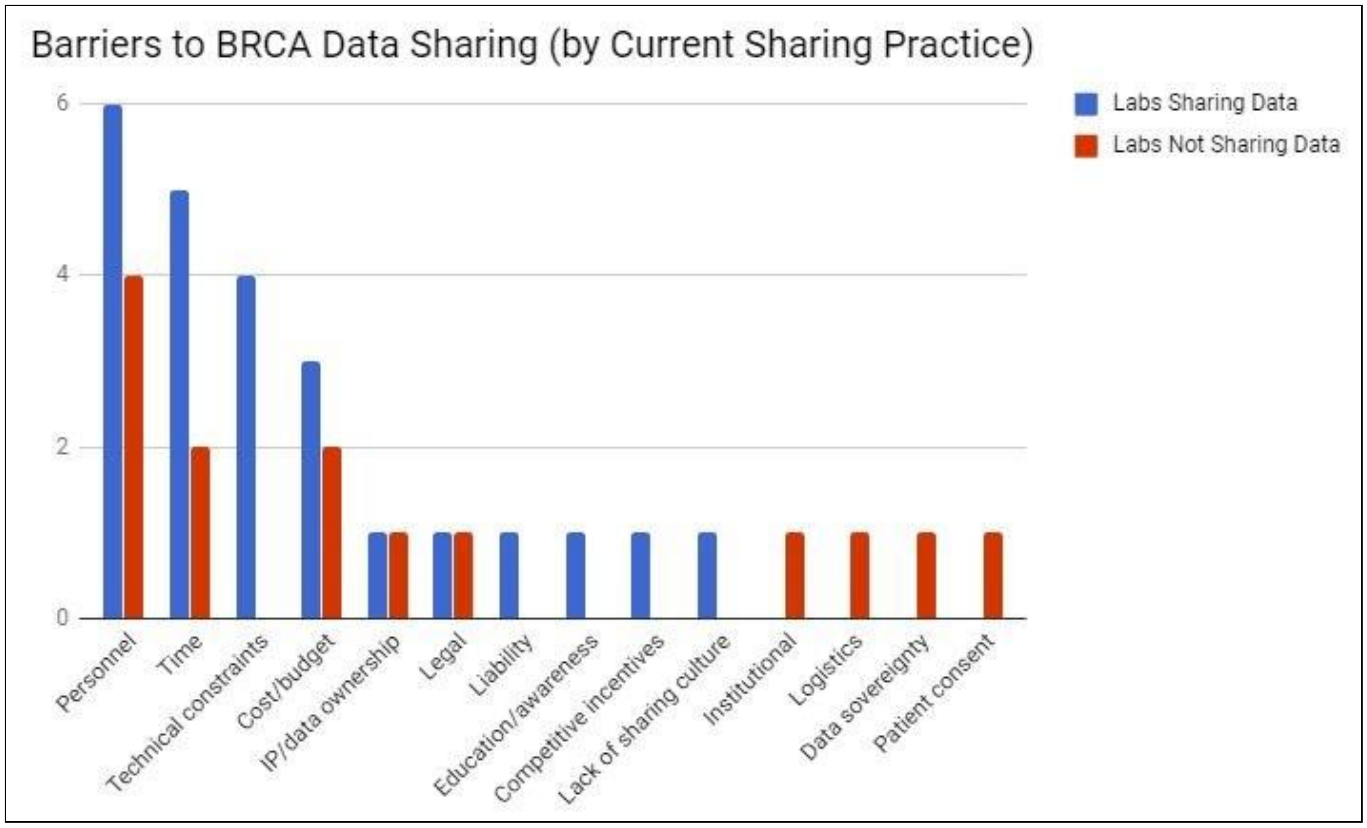


Figure 5. Incentives for BRCA data sharing reported among 10 international laboratories.

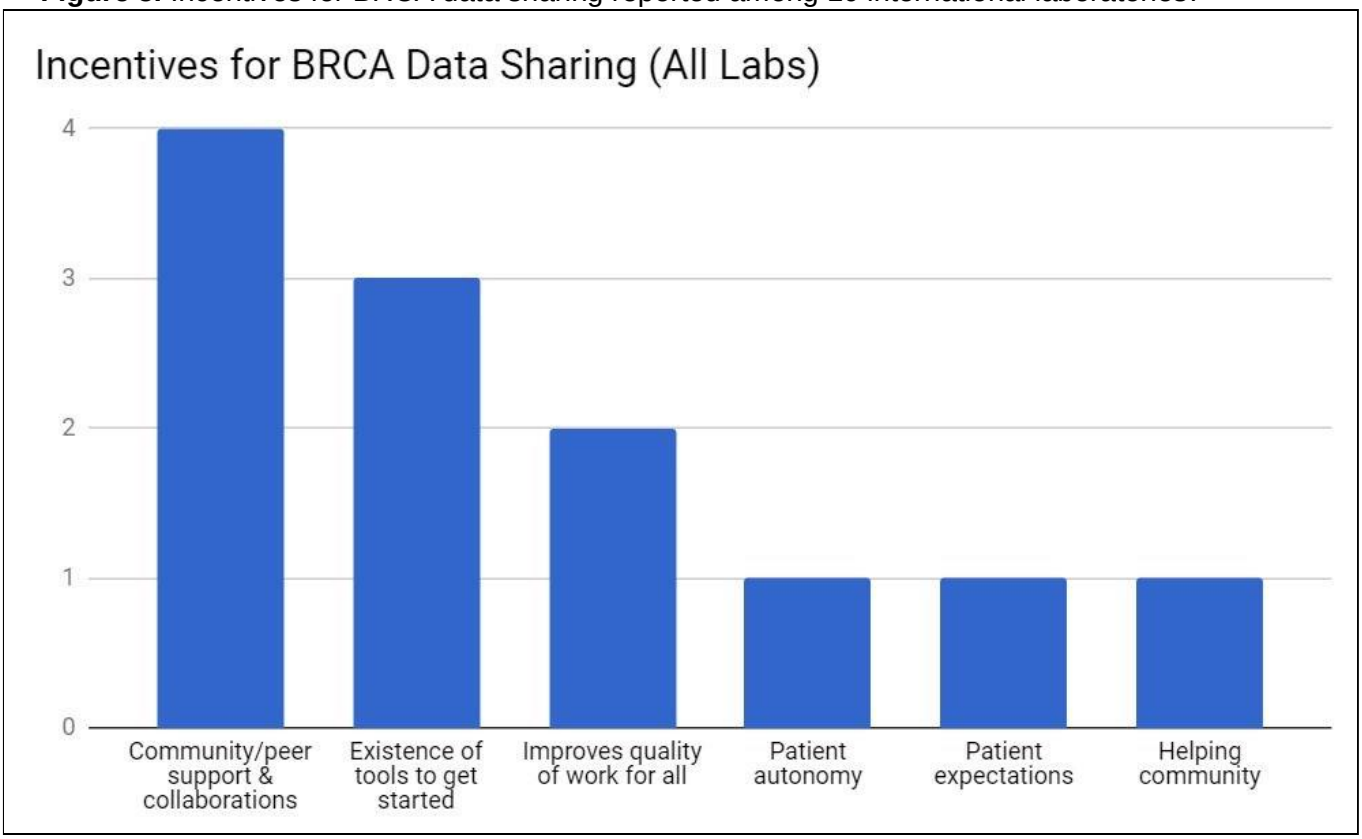


Table 5. A summary of the current barriers to, and incentives/motivations for, BRCA data sharing (as well as additional factors that would further motivate sharing) among 10 international laboratories.

Country	Barriers to Sharing	Current Incentives/ Motivations for Sharing	What Else Would Motivate Sharing?
Labs currently sharing data			
Argentina	Time; Cost/budget; Personnel	Community/peer support & collaborations; Existence of tools to get started	
Brazil	Time; Personnel; Technical constraints; Liability	Helping community; Patient autonomy	
Canada	Time; Cost/budget; Personnel; Technical constraints		Requirement for lab licensing; Data sharing on requisitions / consents
Malaysia	Personnel; Technical constraints	Community/peer support & collaborations; Improves quality of work for all	
Nigeria	Time; Personnel; Technical constraints	Patient expectations	
Tunisia	Time; Cost/budget; IP/data ownership; Personnel; Legal; Education/awareness; Competitive incentives; Lack of sharing culture	Community/peer support & collaborations; Existence of tools to get started	Training to learn increase awareness of sharing databases
Labs not currently sharing data			
Mexico *(plans to share)*	Institutional; Legal; Personnel		
Qatar *(plans to share)*	Time; Cost/budget; Personnel; Logistics		Policy changes; Resources
South Africa *(plans to share)*	Personnel; Data Sovereignty; Patient consent	Community/peer support & collaborations; Improves quality of work for all (best practices)	Existence of tools to get started
Turkey *(open to sharing in future)*	Personnel; Cost/budget; IP/data ownership; Time		Dedicated personnel; Opportunity to receive financial gain for sharing; Existence of easy, reciprocal, curated database for sharing

Table 6. Detailed responses relating to BRCA data sharing barriers and incentives/motivations, from among 10 international laboratories.

Further Comments Relating to Data Sharing Barriers
Cost/Budget ● There are no grants to support data sharing efforts *Technical & Time Constraints* ● Sharing with ClinVar takes time (data preparation), has technical constraints (no automated submission, requires manual creation & review of Excel spreadsheets) ● Lack of local databases & tools adapted to Tunisian and North African populations (e.g. family history tools, risk prediction algorithms) makes it difficult to accurately interpret variants for our patients; this slows down data generation & interpretation rate, which thus affects ability to share ● Sharing with COGR requires converting variant data into a specific format ● Requirement to apply ACMG classes before sharing with some databases is a constraint (ACMG classes don't work well for moderate risk alleles) ● Lack of standards and automation to support sharing ● Labour-intensive, "extra task" on top of regular work; a "thankless job" *Institutional* ● We work for a private health system in which data sharing is not possible (at present). Initiation of sharing will require institutional permission from IRB, Board of Directors, and national Health Regulatory Agency *Legal & Liability Issues* ● Sharing interpretations exposes lab to liability ● Currently illegal in Tunisia (and many other African countries) to share genetic and epidemiological data internationally, but we are in discussions with policy makers to attempt to change this *IP & Data Ownership* ● If our shared data is used by others to produce intellectual materials, what are our rights/protections in terms of IP? If our data is shared with third parties for financial gain, what is our share of that profit? What if a database starts selling our data to drug companies? It's our data to start with *Data Sovereignty* ● In African context, concerns around exploitation and misappropriation put a high responsibility on those choosing to share data; makes robust consent very important *Patient Consent* ● Challenges in obtaining patient consent to share data include low access to trained personnel / genetic counselling services to explain these concepts, low literacy rates in some patient populations, and concerns about insurance discrimination *Lack of Sharing Culture* ● Need to promote a 'spirit' of collaboration and data sharing across Africa ● There is need for an international collaboration with groups/consortia working on BRCA variants to learn more about their experience in data sharing and to further our BRCA interpretations
Further Comments Relating to Current or Theoretical Data Sharing Incentives

Community & Culture

- Data sharing is the best thing that can happen for people working in genetic testing. Sharing is the basis for biological meaning; finding a variant that could be clinically useful worldwide is the best reward in your work
- Sharing should be reciprocal; those benefiting from use of a database should contribute as well

Existence of Tools to Get Started

- Following examples of other consortia and their organizational models helped us to establish our own data-sharing processes
- We are seeking guidance on models for sharing: want to follow processes/guidelines that are internationally-endorsed and robust (e.g. must all variants be Sanger-confirmed?)

Education & Awareness

- Need training opportunities to increase awareness of existing/emerging databases for data sharing

Database Improvements

- Would be further motivated to share with ClinVar if there were periodic evaluations/curations to clean up data (e.g. if a consortium reviewed all novel variants on a regular basis and gave consensus classifications to reduce noise on site)

Lab Licensing

- If lab licensing required data sharing, this would be major incentive; would force labs to hire dedicated personnel for this task

Patient Autonomy

- Strongly believe that patients have a full right to all data relating to them; we will send patients any of their data that they request

Patient Expectations

- Patients want us to share the data

Financial Gain

- An opportunity to receive financial gain for sharing would be an incentive

Databases

This section includes the summary results from among eight international databases in seven countries interviewed (in Brazil, Canada, France, Germany, Japan, Netherlands, and United Kingdom).

Experience Hosting BRCA Data in a Shareable/Accessible Context

The interviewed databases (Table 7) are each hosted in a way that enables sharing in an online context, at least among consortium members and many allowing fully open access. The earliest web-based data release of BRCA variant data was by what has now become Leiden Open Variant Database (LOVD, Netherlands) - in promotion of data access and variant classification, a list of variants was made available online in 1995. Others began maintaining shared spreadsheets of variants among several groups in the same timeframe and launched online access later, and other groups have launched databases more recently, in the past few years. Databases that launched in the 1990's have more total variants than those which have launched more recently. Access methods vary across the databases, with some allowing for full access, some requiring registration, and others being closed except to participating national consortium members. Databases also provide variable levels of curation, including none at all, submitter-based, or relying on specific curators.

Table 7. Summary of findings across eight interviewed databases.

Country hosting database	Year developed	Total unique BRCA variants	Access	Curation methods
Brazil	2017	180	Variants: open Phenotypes: under discussion	Submitted by labs performing the testing
Canada	2012	2366	Reciprocal access for labs that share with them	Submitted by labs performing the testing
France	1995 (spreadsheet) 2012 (database)	5955	Registration required Free for academics License for companies	French consortium VUS classification group
Germany	1996 (consortium) 2006 (web portal)	3500	Reciprocal for consortium members only	German consortium classification task force for VUS
Japan	2016	500	Subscription required for approved users	Look to other databases to support curation

Netherlands	1995 (list of variants online)	8058	Submitters decide which data are made public	Assign a curator for each gene; accepts classifications from submitters
United Kingdom	2014	35,000 total (not unique) records	Developed a 'pseudonomization' software to share genetic information with other collaborators	No curation provided
United Kingdom	2013	22 open access patients with variants	Fully open for 25k consented cases	No curation provided

BRCA Data Access and Download Methods

The eight databases were asked about their data access methods and whether any data stored were available to other users for download (Figures 6 and 7).

Databases have been developed to allow a variety of access methods, depending on the desired end-user community. Six databases were developed around regional data collection. Three communities have set up specific national consortia to aggregate and interpret the variants identified in testing laboratories around each country. France's Database is also available to view in a portal, though not for download. One national community in the UK, has aggregated longitudinal public health information to support clinical use in the UK, but has also developed a specific 'pseudonomization' software to enable sharing certain data, including genetic data, more broadly while preserving privacy and security for tested individuals. Canada's database is populated by a consortium of Canadian and American laboratories to develop technologies to support medical decision-making, including variant classifications, and allows variants and classifications to be downloaded. Brazil's database similarly has arisen around a consortium of Latin American testing laboratories and national data aggregation efforts, and allows variant-level data downloads.

Two databases have developed into international repositories, and in one case, as a specific database and data-sharing software that can be used by any group to store and share data. Both of these resources share variant- and case-level data for appropriately consented cases, and are strong promoters of data sharing to improve diagnosis and care for individuals with genetic diseases.

Figure 6. Access methods for variant- and case-level databases among databases.

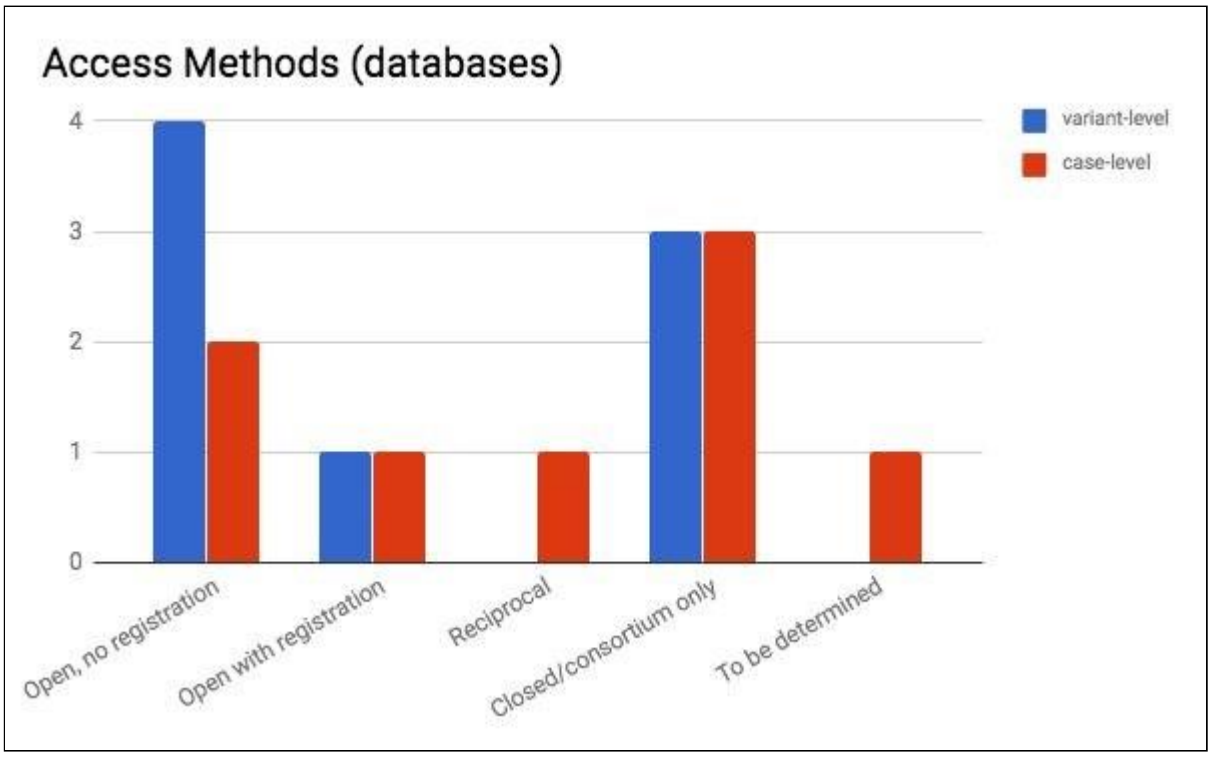
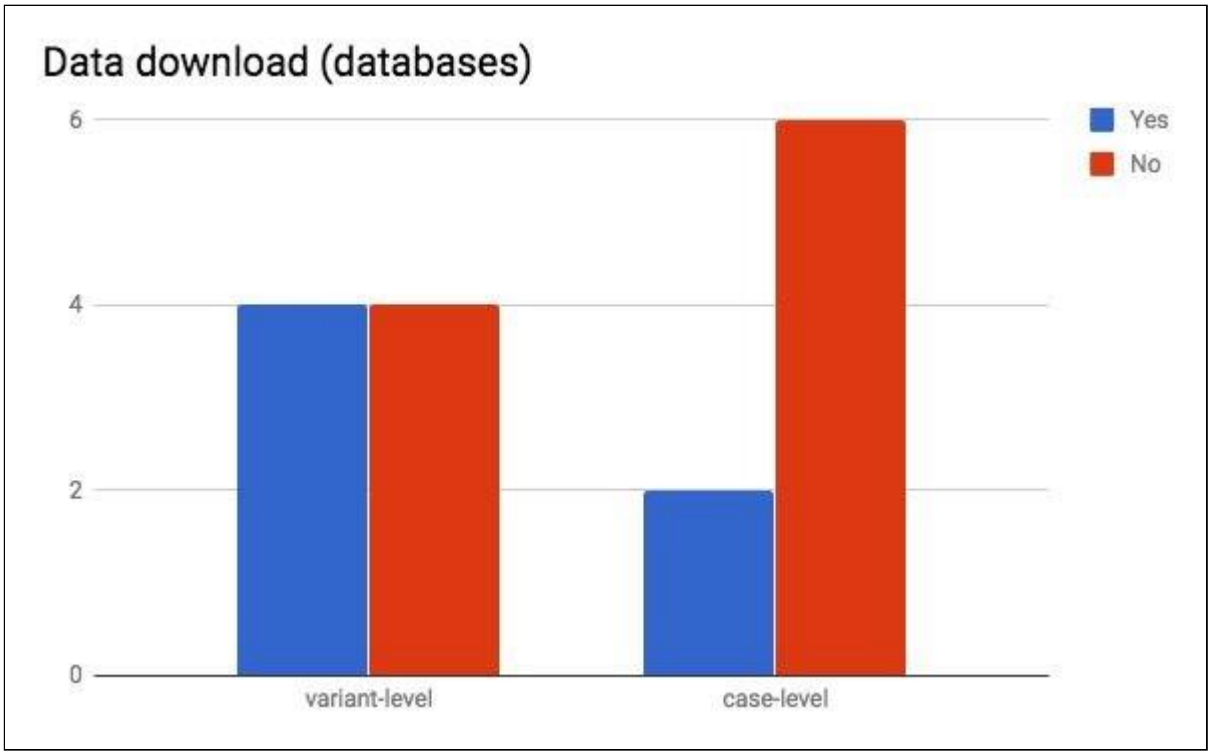


Figure 7. Practices of databases to permit data downloads.



BRCA Data Sharing: Current Practices, Barriers, and Incentives

All databases were developed in order to promote data sharing for improved variant classification and clinical decision-making. However, databases have made variable decisions about with whom to share, as well as which information to share, when sharing information beyond the original users of the shared data (Table 8). These decisions have been made by balancing the diverse barriers to and motivations for data sharing experienced by each database (Figure 8 and 9; Table 9).

Among the eight international databases interviewed, six currently share their data outside of the database, with the two exceptions being in Japan and the UK. Both of these resources are actively pursuing strategies to share and plan to begin sharing shortly, including substantial interest in sharing variant data with open resources such as BRCA Exchange, and in one case the development of an in-house pseudonomization software to enable sharing of information while maintaining privacy for individuals.

Five databases (in Brazil, Canada, France, Netherlands, and UK) allow all users to view appropriately consented variant information, and those in the Netherlands and UK also allow viewing of case-level information that has been appropriately consented. As mentioned above, four of these resources also allow for download of information (Figure 7). The German database does not enable public viewing of its data, but it does share with specific variant interpretation consortia such as CIMBA and ENIGMA, and does allow ENIGMA to share expert classifications of German variants with other resources such as ClinVar and BRCA Exchange.

Table 8. Summary of BRCA data sharing practices among databases

Country hosting database	Whom Sharing With	Data Elements Shared	Personnel Involved in Sharing	Consent Required to Share in Database?
Brazil	All users can view variants and classifications	variants; phenotypes under discussion	one FT staff, many PT including students, postdocs	Yes
Canada	All users can view variants; member labs have access to case-level data; ClinVar receives variants	variants and classifications viewable	one FT staff, many volunteers	No
France	All users (variant classifications); CIMBA; ENIGMA	variants, classifications, case level data, functional data, co-segregation data	Some staff; many volunteers	Yes
Germany	CIMBA, ENIGMA	variants, family history, in vitro assay results	Consortium members volunteer their time	Yes
Japan	n/a	n/a	n/a	n/a

Netherlands	All users can view data that have been consented to be shared	variants, classifications, case level data if submitter makes available (opt in)	one FT staff	No
United Kingdom	Working out how to share with ENIGMA and BRCA Exchange using new software	phenotypes and genetic info when approved to use pseudonomization software	Office for Data Release negotiates data released	Yes
United Kingdom	All users can view data that have been consented to be shared	variants and deep phenotypes	Some staff, many volunteers	Yes

When asked about barriers to BRCA data sharing, five databases mentioned a lack of time as a key barrier. Five databases also mentioned issues related to data ownership, such as cultural expectations that submitters 'own' submitted data and must agree for it to be shared. Other barriers mentioned included cost, legal barriers, and the lack of a sharing culture, which often results in databases not sharing in order to maintain public trust. Logistics (e.g., challenges around how to ensure an 'opt-out' option for individuals), personnel, technical constraints, and patient consent were also mentioned as barriers.

The databases have all been developed with sharing in mind, so it comes as no surprise that the majority of databases specifically mentioned the strong motivation to share derived from the mutual benefit provided to other testing groups in order to improve the quality of work for all. A similar common motivation is to help the community, including clinicians and individuals receiving care, make more informed medical decisions. In some cases, technology has also proven to be a motivation. For instance, the Canadian database is motivated to share because there are existing software resources to support sharing. The Netherlands database itself specifically incentivizes sharing through the development of its own open-source software which allows users to share any data they choose, and allows microattribution for individual submitters as a further incentive for sharing with a central repository. Other incentives include collaborations with other members of the scientific community and patient expectations for data to be used effectively for better research and health outcomes.

The breakdown of responses per database is recorded in Table 9.

Figure 8. Barriers to BRCA data sharing reported among eight international databases.

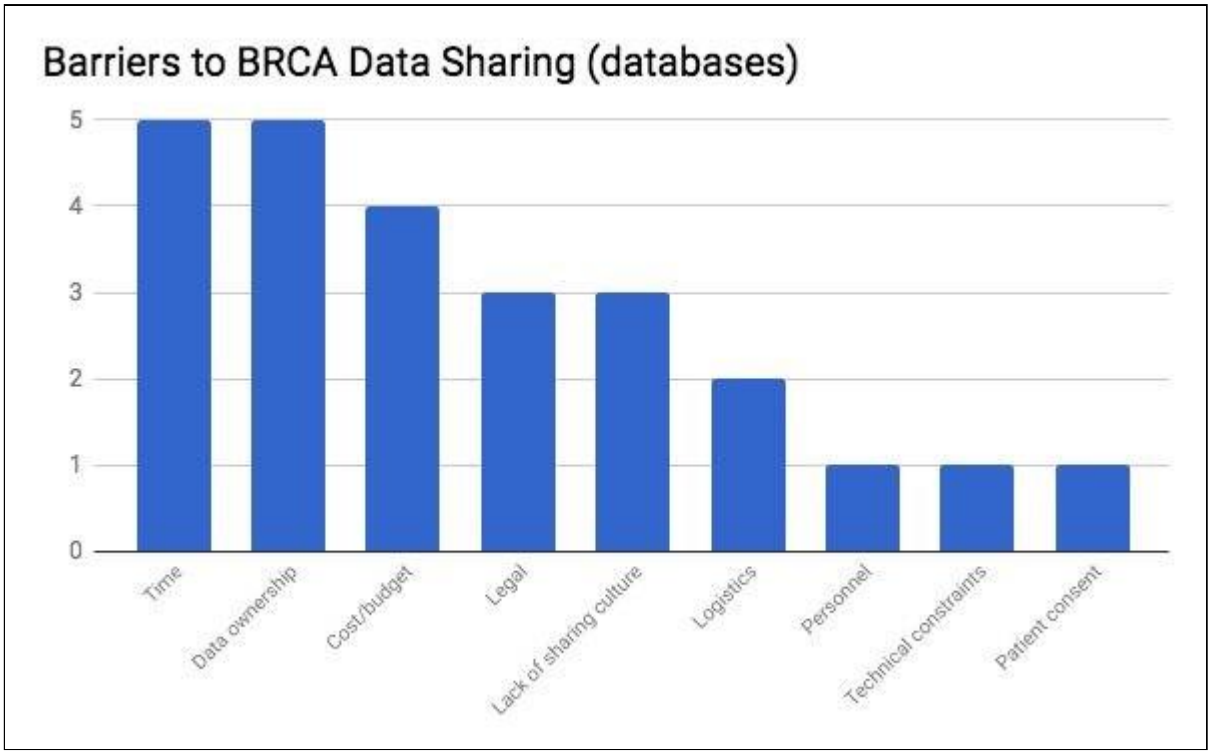


Figure 9. Incentives for BRCA data sharing reported among eight international databases.

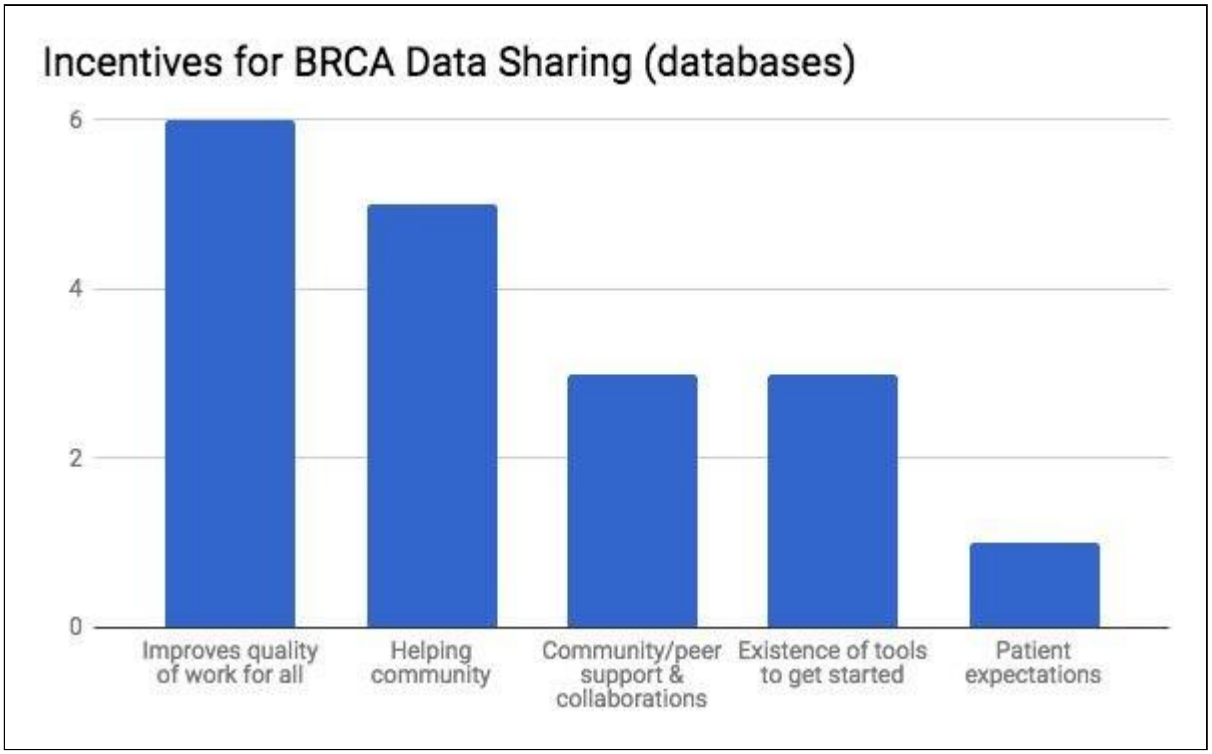


Table 9. A summary of the current barriers to, and incentives/motivations for, BRCA data sharing (as well as additional factors that would further motivate sharing) among eight international databases.

Country hosting database	Barriers to Sharing	Current Incentives/ Motivations for Sharing	What Else Would Motivate Sharing?
Brazil	Convincing commercial labs to share; Consent	Examples from community including GA4GH	
Canada	Time; Cost	Community; Good existing software; Precedent for sharing	
France	Time; Cost; International regulations	Enabling use of high quality data by anyone who needs it	A federation of databases to ensure sustainability
Germany	Time for personnel to participate in consortium; Time for decision making by consortium; Culture around submitter ownership of data	Task force for VUS to enable more variant classification	Ability for consortium members to maintain oversight over their submitted/shared data
Japan	Unsure how to start; Hard to achieve consensus among consortium members		
Netherlands	Time; Cost; Each submitter has to agree to share; No method to share bulk data (laborious, manual process)	Microattribution; Software (can download and label with own logo)	
United Kingdom	Time; Cost; IRB approval; EU data governance and regulations; Opt-out process required; Concerns about privacy - need to maintain public trust	Benefit public health	Software to enable privacy while sharing
United Kingdom	Governance; Culture (wariness toward sharing); Sense of ownership of data by submitters	Help individuals in rare disease families find causal genes/variants	

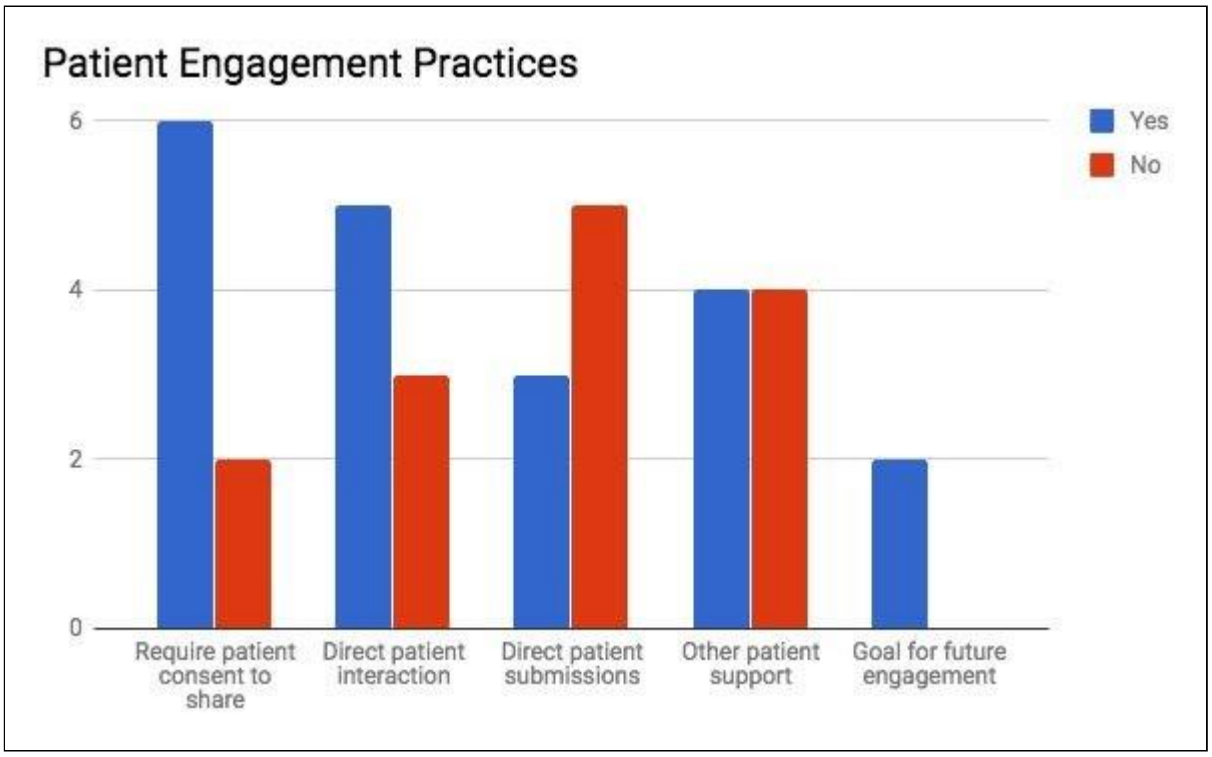
Database practices for patient engagement and consent

Genetic databases rely on the submission of patient information to develop resources that can be used for classification and medical decision-making by the clinical and research communities, but have highly variable engagement strategies with patients and participants themselves (Figure 10 and Table 10).

All databases were launched with an aim to benefit the clinical community and its ability to inform medical decisions among members of the population who discover that they carry a BRCA1/2 variant that increases their risk of developing cancer. However, more and more, individuals themselves are becoming interested to use and contribute their own data and that of the larger

participant and variant carrier communities. Among the databases, six resources require explicit patient consent in order to share participant information. (Consent is attained at the point of care by the individual's clinician.) Beyond that, five databases support direct patient interactions, three support direct patient data submissions, and four have other patient support services. Two databases expressed the goal to develop patient engagement strategies in the future.

Figure 10. Patient engagement and consent practices among eight international databases.



Conclusion

Overall, our interviews found that robust BRCA1/2 data sharing is already taking place in both testing laboratories and databases around the world. There are substantive barriers preventing sharing in some contexts, but there is also a general consensus for the value of sharing to support both variant classification and well-informed clinical decision-making among providers and variant carriers. With the right mixture of legal and regulatory support, technological development, and financial reimbursement, it is our belief that the world's data could be accessed to an even great extent than now, for the further improvement of medical understanding and support of BRCA1/2 variant carriers and their families.