

Association for Molecular Pathology Position Statement: Variant Data Sharing

July 29, 2021

Background

The sharing of data, samples, and other health information is critical for advancing healthcare. It is essential both for understanding the contribution of genetic and genomic variation to disease and conditions, and for translating that information through the development, validation, and interpretation of clinical testing. Moreover, it contributes to the research and development of life-changing vaccines and therapies. Facilitating the sharing of genetic variants will have far-reaching effects – from individuals working to elucidate the contribution of certain human sequence variants of rare/orphan disorders to those working in the field of SARS-CoV-2 diagnostic testing and variant tracking. Making data broadly accessible reduces fragmentation of knowledge from multiple small studies and accelerates the ability of medical professionals to determine the etiology of a patient’s condition. In addition, it supports innovation while also encouraging a standardized approach to providing high quality care whenever possible.

One area where data sharing has proven extremely effective is the categorization of variants identified through genomic sequencing of individuals.¹ The National Institutes of Health (NIH) recognized the importance of the sharing of data associated with human genetic variation when they created ClinVar, a freely accessible, public archive of genetic variants and their classifications with supporting evidence and associated phenotypes in many cases.² Submissions of de-identified data to this curated database accelerates the process for re-assignment of variants of unknown significance (VUS) to clinically actionable categories (*e.g.*, benign or pathogenic), which NIH considers a critical aspect of the quality assurance process for accurate genetic and genomic testing. Additionally, NIH’s ClinGen (clinical genome) resource provides an avenue for organizations to be recognized by ClinVar as an expert panel or provider of practice guidelines.³

In 2017, the American College of Medical Genetics and Genomics (ACMG) published a position statement calling for the extensive sharing of laboratory and clinical data from genetic testing.⁴ Their position statement on DNA-based screening and population health published in 2021 calls for

¹ https://clinicalgenome.org/docs/?doc-type=publications#list_documentation_table

² <https://www.ncbi.nlm.nih.gov/clinvar/intro/>

³ <https://www.ncbi.nlm.nih.gov/clinvar/docs/clingen/>

⁴ 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 (ACMG). *Genet Med* 19, 721-22 (2017) <https://doi.org/10.1038/gim.2016.196>

organizations, both public and private, to share de-identified variant data and accompanying evidence supporting their clinical impact plus health outcomes related to genetic penetrance and expressivity.⁵ AMP agrees with ACMG that broad data sharing is necessary to improve patient care and to advance the development of tests and treatments. An example of how shared genetic data improved patient health was the elucidation of VUS in hereditary breast cancer. Prior to the 2013 Supreme Court decision in *Association for Molecular Pathology vs. Myriad Genetics Inc.*⁶ much of the *BRCA1* and *BRCA2* variant data obtained through hereditary breast cancer testing was considered proprietary and inaccessible to molecular professionals working to improve the care of their patients. Through the creation of an open access database, the Breast Cancer Information Core, molecular professionals were able to overcome this barrier and more easily resolve a VUS in these genes than they were prior to the court's decision.⁷ The information in that database is now incorporated into ClinVar.

Despite the current availability of public databases, multiple recommendations, and calls to action, as well as some incentives to contribute to them, participation in data sharing among clinical laboratories continues to remain low. ClinGen reports that as of January 2020 only fifteen clinical laboratories within the United States met their minimum requirements (e.g., 95% of reportable sequencing variants with supporting evidence for pathogenic/likely pathogenic variants from past two years) for data sharing to be recognized in this program.⁸ Thus, information that is vital to the advancement of molecular pathology and improvement of patient care continues to be siloed.

Additionally, there is a lack of population diversity in the genetics research literature and most existing clinical genetic variant databases.⁹ In 2021, the American Heart Association (AHA) published a position statement supporting three principles for ethical sharing of genomic data from Indigenous communities and marginalized racial and ethnic groups for research – building trust, enhancing accountability, and improving equity.¹⁰ AMP agrees with the AHA that “respect, honesty, justice and fairness, reciprocity or assurance of mutual benefit, and care for the individual and community” are crucial for the sharing of genetic variant data for research and that similar principles should be considered for the sharing of clinical variant data. The current dearth of population-specific variant data in most commonly utilized databases is a disservice to many patients. To remedy this inequity, there should be a strong push to encourage clinical laboratories to submit variant data for patients and normal controls from all ancestry backgrounds, including marginalized racial and ethnic groups and Indigenous peoples, to easily-accessible databases.

⁵ Murray, M.F., Giovanni, M.A., Doyle, D.L. *et al.* DNA-based screening and population health: a points to consider statement for programs and sponsoring organizations from the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 23, 989-995 (2021). <https://doi.org/10.1038/s41436-020-01082-w>

⁶ Assoc. for Molecular Pathology v. Myriad Genetics, Inc., 569 U.S. 576 (2013) https://www.supremecourt.gov/opinions/12pdf/12-398_1b7d.pdf

⁷ Toland, A.E., Brody, L.C. & the BIC Steering Committee. Lessons learned from two decades of *BRCA1* and *BRCA2* genetic testing: the evolution of data sharing and variant classification. *Genet Med* 21, 1476–1480 (2019). <https://doi.org/10.1038/s41436-018-0370-4>

⁸ <https://www.clinicalgenome.org/tools/clinical-lab-data-sharing-list/>

⁹ There are a few example databases incorporating information from a more diverse population including the Genome Aggregation Database (gnomAD).

¹⁰ Mudd-Martin, G., Cirino, A.L., Barcelona, V., Fox, K., Hudson, M., Sun, Y.V., Taylor, J.Y., Cameron, V.A., and on behalf of the American Heart Association Council on Genomic and Precision Medicine, and Council on Cardiovascular and Stroke Nursing; and Council on Clinical Cardiology. Considerations for Cardiovascular Genetic and Genomic Research With Marginalized Racial and Ethnic Groups and Indigenous Peoples: A Scientific Statement From the American Heart Association. *Circ Genom Precis Med*; 14:e000084 (2021).

To better understand why clinical laboratories might be unable to routinely submit variant data, the Association for Molecular Pathology (AMP) convened a working group to explore the barriers to participation and to put forth recommendations to encourage all relevant stakeholders to promote the engagement of these practices and facilitate ease of access to this important information. While variant data sharing across all molecular pathology sub-specialties (*i.e.*, genetics, oncology, and infectious diseases) is vital, additional distinct considerations for sharing exist within each sub-specialty. This is especially true for infectious diseases and the sharing of pathogen variant data. For example, standardization of nomenclature and classification of pathogen variants are inconsistent (*e.g.*, SARS CoV-2 B.1.1.7. vs. alpha variant) and have the potential to lead to multiple pathogen-specific databases where each database has different and reasonable submission requirements for a specific pathogen. Thus, the recommendations made in this statement should be considered a starting place for the sharing of variant data, broadly, and should be incorporated with discrete and/or additional needs for each sub-specialty.

Barriers to Variant Data Sharing

Several factors may contribute to a laboratory being unable or unwilling to either submit sequence variant information or to utilize information stored in an existing database. AMP's working group identified the following practical challenges for both accessing and participating in a curated variant database.

Limited Resources:

Data sharing has a cost in time, resources, and effort that require support. The processes needed to prepare for sharing data – for example de-identifying a patient's health information before submission to a variant database; ensuring that the data meets at least the minimum requirements expected from ClinVar or other data sharing entities; and standardizing the format of the data to send it to the online database – are highly variable and labor intensive. Clinical laboratories already experience insufficient resources in terms of time, personnel, or communication channels required for sharing variant data with public databases. Additionally, laboratories face challenges of insufficient reimbursement, and sharing variant data is not a reimbursable service nor can the expense be passed on to patients. These facets could constitute a disincentive to devote resources for database contributions despite possible benefit to the public.

Ownership of the Data:

Many commercial and noncommercial entities view all health data generated by them, including variants identified in patients' samples, to be proprietary and hence, opt to monetize the data and/or restrict its use for only the entity's benefit. Often, employees of those organizations may be instructed not to share any data, including de-identified variant and phenotypic information.

Assurance in Database Curation:

Genetic variant data identified by clinical laboratories are derived from a variety of medical disorders including inherited conditions, cancer, and infectious diseases. Additionally, curation practices for the associated molecular pathology sub-specialties vary among the professional communities, thus, some end-users of the data might have varying degrees of confidence about the accuracy of variant data and databases. For example, some somatic variant databases currently lack uniform standards and expert

curation, which can decrease confidence or utility in these databases. Lack of confidence in the accuracy, quality, or completeness of the information in a database, especially as new information is compiled, and reclassified, can lead to hesitancy among clinicians to both contribute to and utilize databases. Professional standards for classification of variants, such as those published by the ACMG and AMP for inherited conditions¹¹ and AMP, American Society of Clinical Oncology (ASCO), and College of American Pathologists (CAP) for somatic conditions¹², may not be uniformly adopted nor implemented. Guidelines also need to be continuously evaluated and updated to standardize and provide an enhanced level of assurance. Moreover, even with successful establishment and adoption of a standardized approach to variant classification, current guidelines may be insufficient for new variant classification emphasizing the essential role of professional judgement for meaningful interpretation.

Protection of Sensitive Information:

The existence of privacy laws, such as the Health Insurance Portability and Accountability Act of 1996 (HIPAA), allow for the sharing of de-identified protected health information (PHI) through mechanisms that protect patients. Despite this, some concerns may persist about data privacy and data security, which may make some organizations, institutions, or companies hesitant to participate in sharing practices. Efforts that work to provide clear guidance on sharing of de-identified PHI would help to minimize hesitancy. For example, while not necessary, it would be ideal to obtain broad consent from patients for variant sharing during the specimen collection process.

AMP Position on Variant Data Sharing

As the COVID-19 pandemic response has demonstrated, laboratories, physicians, manufacturers, patients, and researchers benefit from shared data resulting from diagnostic testing and genome sequencing. For instance, improved and more frequently generated sequence data of the SARS-CoV-2 viral genomic variants shared throughout the medical community has improved patient care practices, infection monitoring, biotech developments, diagnostic testing, and population surveillance. Using a similar approach, more universal sharing of human genomics sequence data for somatic and inherited disorders is expected to improve diagnosis and clinical management of affected patients, particularly for rare disorders. More robust variant data sharing may also address health inequities by improving understanding of the types and frequencies of variants in different subpopulations. While databases for the sharing of genetic information about these conditions are frequently used by research laboratories (and may be a requirement of their research funding), clinical genetics laboratories' contribution to these resources may not be widespread. The benefits of data sharing are obvious and many research and clinical laboratories benefit from what others are already publicly contributing, but advancement in the field and patient care is limited by the low number of laboratories that participate in this critical information exchange.

¹¹ Richards, S., Aziz, N., Bale, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17, 405–423 (2015). <https://doi.org/10.1038/gim.2015.30>

¹² Li, M.M., Datto, M., Duncavage, E.J., *et al.* Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn* 19, 4-23 (2017) <http://dx.doi.org/10.1016/j.jmoldx.2016.10.002>

The Association for Molecular Pathology calls upon all relevant individuals and organizations, including hospitals, academic medical centers, commercial diagnostic laboratories, patient organizations, policymakers, and others, to support and facilitate the sharing of molecular genetic variant data and offers the following recommendations.

Recommendations for organizations, institutions, and companies:

- All relevant organizations, institutions, and companies should publicly endorse the value and support database contribution of variant data for the collective benefit of current and future patient care. Support should include efforts to provide guidance on sharing of de-identified PHI.
- These stakeholders should implement data sharing policies that allow treating clinicians, laboratory professionals, and researchers to transfer de-identified variant and phenotypic information to publicly accessible databases such as ClinVar.
 - These data sharing policies should emphasize the importance of contributing data which represents the diversity of their patient populations.

Recommendations for clinical laboratories:

- Clinical laboratories should assign dedicated personnel to interface between shared sequence databases and diagnostic genetic test result platforms. Realizing that a barrier to this may be financial support, clinical laboratories should work with other stakeholders and policymakers to assist in developing additional mechanisms where resources and staff could be better supported.
- Clinical laboratories should gain better understanding of the clinical relevance of variants in different populations, especially any that are particular to a given laboratory's clinical catchment. This will lead to enhanced diversity within variant databases and help reduce existing barriers to test utilization in underserved populations.
- Clinical laboratories contributing to databases should be acknowledged publicly to establish data sharing as the norm. To this end, laboratories should leverage public communications, websites, etc., to relay their own support of these contributions.
- Clinical laboratories should strive to use harmonized/standardized test requisition forms for common data elements. This will help to standardize the submission process, enhance the ability to compare data across laboratory submissions to variant databases, and better aggregate data from patients across laboratories.

Recommendations for operators of public databases:

- Operators of public databases should strive to establish convenient and practical data submission processes to reduce the burden of data sharing. Ideally, this should involve input from clinical laboratories to identify and address common barriers. Database managers and curators should provide support for resource-strained laboratories who want to contribute.

- To ensure user confidence, operators should also work to apply curation processes to remove obvious data outliers; request evidence when it was not included in the original submission; and follow up with individuals who have submitted older data to verify relevance (*e.g.*, that new VUSs/likely pathogenic/likely benign variants have been updated).

Recommendations for policymakers:

- Relevant US and international government agencies working with clinical, academic, and public interest stakeholders should convene a standards-setting body to accelerate harmonization of variant database elements, including:
 - Establishing common data elements that capture information on variables collected by clinical laboratories for submission purposes. This would work to improve interoperability and accessibility of genetic test results within electronic health systems.
 - Harmonizing sequence variant classification schemes with integration of pathogenicity criteria, including for demonstrating any interrelation between germline and somatic variants.
 - Establishing standards for the collection of clinical data on genetic testing requisition forms and electronic systems whenever possible for a disease or condition.
- Policymakers should consider incentives for variant data sharing including, but not limited to, enhanced reimbursement to offset the costs of data submission or additional opportunities for public recognition.
- Policymakers should ensure that intellectual property policies are maintained to support technology development around genetic testing without jeopardizing variant sharing needs for the broader clinical community by:
 - Taking action to restrict ownership of collected genetic information by entities.
 - Preventing changes to the US patent system that would allow the patenting of biomarkers, such as genes, and their association with health status.
- Policymakers should seek to address barriers for inclusion of underserved populations both in public sequencing databases and in clinical test utilization. Remedies to such barriers should be mandated in any future healthcare, clinical laboratory, or research funding policy reform efforts.

Recommendations for patients, providers, and researchers:

- Patient groups, professional societies, and other stakeholders representing specific disease areas should work to harmonize phenotypic data collection that are essential for variant classification.
- Patient groups devoted to health initiatives for underserved populations should work with providers and laboratories to increase diversity within databases, especially through the identification and remedy of barriers, *e.g.*, socioeconomic or cultural, that impact access to sequencing.
- These stakeholders should explore ways that information gleaned in a research setting can be more quickly translated for use in clinically focused databases.

Conclusion

Public sharing of variant data across clinical laboratories is a valuable mechanism for establishing important tools and resources to further patient health and the field of precision medicine. AMP believes that adopting these recommendations will assist in increasing the use and standardization of variant databases, which will improve their ability to meaningfully contribute to the work of genetics professionals to positively impact patient care. Thus, AMP is strongly committed to working collaboratively with all relevant stakeholders to meet these recommendations with the goal of establishing data sharing as the expectation and norm.

As we look beyond implementing the recommendations, it is important to consider that today, patient level clinical data is an optional data element in variant databases. More attention will have to be given to working towards interoperability between variant databases and electronic health systems so that genetic information can be better incorporated into professional practice and patient care. Moreover, the contribution of information between variant databases and electronic health systems will facilitate genetic and genomic discoveries into healthcare advances for patients. AMP will continue to provide and update recommendations and guidelines to help achieve this goal.

About AMP

The Association for Molecular Pathology (AMP) is an international medical and professional association representing approximately 2,500 physicians, doctoral scientists, and medical technologists who perform or are involved with laboratory testing based on knowledge derived from molecular biology, genetics, and genomics. Membership includes professionals from the government, academic medicine, private and hospital-based clinical laboratories, and the in vitro diagnostic industry. Through the work of our subject matter experts, AMP continues to develop and update our evidence-based guidelines to foster and support innovation while establishing clinical best practice recommendations.