# A new era in the interpretation of human genomic variation

**Heidi L. Rehm, PhD, FACMG**

Laboratory for Molecular Medicine, Partners Healthcare Personalized Medicine, Cambridge, MA

Department of Pathology, Brigham & Women's Hospital and Harvard Medical School, Boston, MA

The Broad Institute of MIT and Harvard, Cambridge, MA

Just five years ago the world of variant interpretation was very different. There was no standard terminology for variant classification with laboratories using many different terms as ambiguous as "mutation" and "polymorphism" to convey pathogenicity, or lack thereof, and a plethora of qualifiers such as "possibly", "probably", and "likely" to convey degrees of uncertainty. The only terminology broadly agreed upon was "variant of uncertain significance" or VUS. Furthermore, there was substantial variability in how labs evaluated evidence. A single publication in a peer-reviewed journal could be cited as justification for declaring pathogenicity, even though on many occasions little evidence to support that assertion was present in the publication beyond, for example, a non-statistically significant observation in an affected individual. This was perpetuated by the routine deposition of "deleterious mutations" into the Human Gene Mutation Database (HGMD)[1], often without strong evidence supporting these claims. Similarly, exemplar variants included in the OMIM (Online Mendelian Inheritance in Man) database[2] to support gene-disease relationships, were assumed to be pathogenic.

## Evolving standards for variant interpretation

To address these issues, guidance was published in 2015 by the American College of Medical Genetics and Genomics (ACMG) and the Association of Molecular Pathology (AMP) to provide more detailed standards for how to weigh specific types of evidence and combine them to reach a classification for variants in Mendelian disease genes.[3] Standardized terms were also published: pathogenic, likely pathogenic, uncertain significance, likely benign and benign. These guidelines have now been widely adopted by laboratories both nationally and internationally.

While the development of detailed standards has brought greater consistency to variant classification, there remains both a degree of expert judgment and a learning curve in the accurate application of these guidelines.[4] The authors of the ACMG/AMP guideline recognized this and acknowledged that the 2015 guideline was a first step towards the goal of more objectively defining pathogenicity. An assumption was that these rules would see both general improvements as well as specific details added for individual genes and

diseases over time.[3] In this month's issue of Genetics in Medicine, Nykamp and colleagues report on a detailed evolution of the ACMG/AMP guideline, transitioning to a semi-quantitative framework with the use of a point-based scoring system and decision tree matrices (Nykamp et al, 2017 - this issue). The ACMG/AMP committee had not decided to pursue a point-based system at the time over concern that assigning points might lead to a false assumption of the quantitative accuracy of the approach given substantial reliance on professional opinion to assign and weight certain types of evidence. However, it was recognized that a semi-quantitative approach would improve the ability to combine both pathogenic and benign evidence as well as allow more gradation in the relative strength of each type of evidence. The careful approach taken by Nykamp and colleagues to systematically dissect the evidence types presented in the guideline, as well as account for many additional subtleties in the types of evidence encountered in variant classification, should be commended. I anticipate that ClinGen's Sequence Variant Interpretation Working group will carefully consider this approach as it works to provide additional guidance to the application of the ACMG/AMP guidelines and support the ongoing evolution of this field.

As a testament to the utility of a standardized approach to variant classification, in this same printed GIM issue, Harrison and colleagues were able to show the successful resolution of 87% of interpretation differences among four major clinical laboratory submitters to ClinVar, with the majority of variants resolved through application of the ACMG/AMP guideline for variant interpretation.[5] A similar rate of success was previously published with an analysis between two laboratories.[6] An additional paper in this issue of GIM, Lebo and colleagues examined variant interpretation among laboratories in the Canadian Open Genetics Repository (COGR) and reported on a consensus-driven process to variant curation through data sharing and application of the ACMG/AMP guideline.(Lebo et al, 2017 - this issue) In this study, the utility of data sharing through the COGR was demonstrated by using *BRCA1/BRCA2* variant classifications from eleven diagnostic laboratories across Canada as a proof of principle. Using a 3-tier model, discordant rates went from 27% to 14% with only 0.9% having medically significant discordance after consensus efforts.

## Developing a community-centric approach to variant interpretation

In addition to the challenges in standardizing variant interpretation, another barrier arose over the last ten years. Laboratories were having difficulty keeping pace with the resources and expertise required to interpret the millions of variants being identified through high throughput sequencing. It became clear that the most efficient mechanism to solving the problem was crowd-sourcing.[7] If each laboratory shared their work of gathering evidence and classifying variants, every lab would benefit. Most importantly, very few data from clinical laboratory testing were routinely published, preventing access to this large and important body of data. After a few attempts to solve this problem, a group of laboratories turned to NCBI to consider launching a database of shared variant interpretations, now called ClinVar.[8] In parallel, these laboratories banded together through a NIH U41 Genomic Resources grant (now part of ClinGen) to support data collection, standardization, and community sharing through a partnership with ClinVar. This effort was in part modeled after the successful International Standards for Cytogenomic Arrays consortium, which brought

cytogenetic laboratories together after the disruptive technology launch of cytogenomic microarrays into mainstream clinical genetics.[9]

Not surprisingly, when laboratories began sharing data in ClinVar, the differences in variant interpretation methods that had been used over the past ten years became clear, as evidenced by discordant variant interpretation records in ClinVar.[10] However, highlighting these differences through a "share and compare" approach allowed laboratories to resolve differences. In the Harrison study, the sharing of unpublished evidence helped resolve differences for 30% of variants. For some variants, the aggregate evidence from multiple laboratories moved variants from uncertain to either pathogenic or benign.

## Is the glass half full or half empty?

When examining ClinVar, two conflicting arguments have been made. One argument discredits ClinVar, arguing that the discordance observed makes the database untrustworthy and a risk to the community.[11,12] The other view highlights ClinVar as a tremendous asset, allowing laboratories to share curated evidence used in interpretation, to identify discordant interpretations so they can be resolved, and to stimulate research on variants classified as uncertain significance.[10,13,14] (Yang et al, 2017 - this issue) And while I am inherently biased in my views as a funded ClinGen investigator and having worked closely with NCBI to guide the development of ClinVar, I think the rapid growth and usage statistics for ClinVar, and the absence of any documented harm whatsoever, bring an objective view that the database has indeed become an invaluable positive asset.

That said, we have more work to do. In this issue of Genetics in Medicine, Yang and colleagues report a detailed analysis of the data in ClinVar and note that ClinVar submissions from non-clinical laboratory sources (of which there are many) are far more likely to cause interpretation discrepancies than submissions from clinical laboratories, which are generally concordant. Also older submissions tend to disagree with newer ones, and low penetrance variants (which usually have limited medical management implications) have more disagreement than high penetrance ones. These factors appear to play a substantial role in those published comparisons that show high discordance among ClinVar entries.[11,12,15] Comparisons that carefully take such obvious factors into account show substantially higher concordance.[14,16,17] (Yang et al 2017 - this issue)

The same factors could affect any ClinVar user. When ClinVar was launched, a star system was developed to distinguish the level of review that each variant had undergone with the goal of aiding users in determining the trustworthiness of the variant interpretations in ClinVar.[10] Unfortunately, while the distinction between 0–2 stars (single laboratory submitters with (1 star) or without (0 star) interpretation criteria provided) versus 3–4 stars (ClinGen-approved expert panels and professional guidelines) is apparent, the distinction across 0 and 1 star submitters has been less clear with a wider range of quality present, as demonstrated by Yang and colleagues. To ameliorate this problem, both Yang and colleagues, as well as ClinGen investigators, have proposed alternatives to the representation of data in ClinVar. However, while the staff at the National Center for Biotechnology Information conduct an enormous effort to ensure the accurate and standardized

representation of each variant and the submitted interpretations, ClinVar is an archive and does not itself pass judgement on the accuracy of the interpretations. As such, ClinGen is now pursuing an alternate representation of ClinVar data that will support a more nuanced assignment of single star submitter status as well as a more facile ability to flag and remove old and outlier interpretations.

## A call to action

As a clinical lab director for fifteen years, I have seen a broad array of quality in molecular diagnostic laboratories. In my own anecdotal experience, I am now seeing a strong correlation between the quality of interpretations and whether the laboratory is a ClinVar submitter. I have also personally watched laboratories dramatically improve their interpretation methods after engaging in active data sharing in ClinVar. I therefore put out a plea to the community to step up and take action to ensure transparency in variant interpretation, allow laboratories to subject their interpretations to peer review by the community and enable aggregation of evidence to improve knowledge over time. Here are some ways stakeholders can help:

1.  Regulatory agencies – Clinical laboratories are required to perform many tasks to ensure the quality of the tests they offer, though most focus on analytical validity, not clinical validity. While quality assurance activities do require dedicated resources, there is little doubt in my mind that submission to ClinVar will have a more positive impact on the accuracy of tests than checking the temperature on PCR machines and the countless other tasks labs perform daily as required by accreditation bodies. It is time for the College of American Pathologists that accredits the majority of genetic testing laboratories to make submission to ClinVar a requirement of quality assurance.

2.  Payers – Payers of healthcare services only reimburse tests that have demonstrated clinical utility. If there is insufficient quality assurance around the accurate interpretation of a test, it will not have utility, plain and simple. If you are a payer, it is time to make reimbursement contingent on laboratory data sharing and strictly enforce the policy. This has already been initiated by at least one payer (https://www.genomeweb.com/informatics/genomic-variant-data-sharing-gains-support-collaboration-seen-key-interpretation).

3.  Healthcare providers – It is the duty of providers to act in the patient's best interest and that means ordering a test from a lab that shares data and helps our community improve knowledge for the benefit of patient care. If you are a hospital or clinician, order your tests from laboratories who share data. Currently, providers can check submitter status on the ClinVar website: https://www.ncbi.nlm.nih.gov/clinvar/docs/submitter_list/ Soon, ClinGen will launch a webpage to assist in checking this status that will take into account the percentage of reported variants being deposited and other factors.

4.  Journals – curating variant evidence from the literature is a time-consuming task and is highly error prone when researchers publish data that is not subject to the data quality checks performed by database curators such as that done for ClinVar.

> If you are an editor of a journal or you review manuscripts for journals, please make sure submission of interpreted variants to ClinVar is a requirement of publication. Supplemental inclusion of metadata using standardized formats (genomic coordinates and HGVS nomenclature) should also be required.

5. Clinical laboratories and researchers – Enough said, it's time to submit your data!

## The end goal

As we continue to explore the best approaches to data sharing and representation of variant evidence, it is important to keep the end goal of variant interpretation in mind, which is to advance medicine and improve patient care. We must recognize that, similar to the rest of the practice of medicine, we will never be completely concordant in our interpretations of variants, which will inevitably continue to rely in part on professional opinion and evolving evidence. Nevertheless, we must ensure that the professional opinion of geneticists is based on a fully shared body of evidence with open and ongoing community review and that we support this endeavor in a cost-effective manner. A colleague once told me that the airline industry does not compete on safety because safety in flight is a fundamental right, highly regulated by federal law. Instead the airlines focus on the added services they can provide to differentiate their businesses. It is time we considered the sharing of variant interpretations a fundamental right of patients to receive accurate clinical care. Laboratories have plenty of other means to differentiate their services.

## Acknowledgements

## References

1. Stenson PD, Mort M, Ball EV, et al. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Hum Genet. 2017.

2. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. Nucleic Acids Res. 2015;43(Database issue):D789–798. [PubMed: 25428349]

3. 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. Genetics in medicine : official journal of the American College of Medical Genetics. 2015;17(5):405–424. [PubMed: 25741868]

4. Amendola LM, Jarvik GP, Leo MC, et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. Am J Hum Genet. 2016;98(6):1067–1076. [PubMed: 27181684]

5. Harrison SM, Dolinsky JS, Knight Johnson AE, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. Genetics in medicine : official journal of the American College of Medical Genetics. 2017.

6. Garber KB, Vincent LM, Alexander JJ, Bean LJ, Bale S, Hegde M. Reassessment of Genomic Sequence Variation to Harmonize Interpretation for Personalized Medicine. Am J Hum Genet. 2016;99(5):1140–1149. [PubMed: 27843123]

7. Bale S, Devisscher M, Van Criekinge W, et al. MutaDATABASE: a centralized and standardized DNA variation database. Nat Biotechnol. 2011;29(2):117–118. [PubMed: 21301433]

8. Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016;44(D1):D862–868. [PubMed: 26582918]

9. Riggs ER, Wain KE, Riethmaier D, et al. Towards a Universal Clinical Genomics Database: the 2012 International Standards for Cytogenomic Arrays Consortium Meeting. Hum Mutat. 2013;34(6):915–919. [PubMed: 23463607]

10. Rehm HL, Berg JS, Brooks LD, et al. ClinGen--the Clinical Genome Resource. N Engl J Med. 2015;372(23):2235–2242. [PubMed: 26014595]

11. Gradishar W, Johnson K, Brown K, Mundt E, Manley S. Clinical Variant Classification: A Comparison of Public Databases and a Commercial Testing Laboratory. Oncologist. 2017.

12. Vail PJ, Morris B, van Kan A, et al. Comparison of locus-specific databases for BRCA1 and BRCA2 variants reveals disparity in variant classification within and among databases. J Community Genet. 2015;6(4):351–359. [PubMed: 25782689]

13. Harrison SM, Riggs ER, Maglott DR, et al. Using ClinVar as a Resource to Support Variant Interpretation. Curr Protoc Hum Genet. 2016;89:8 16 11–18 16 23.

14. Nussbaum SEL, Shan Y, Melissa SC, et al. Consistency of BRCA1 and BRCA2 Variant Classifications Among Clinical Diagnostic Laboratories. 2017.

15. Balmana J, Digiovanni L, Gaddam P, et al. Conflicting Interpretation of Genetic Variants and Cancer Risk by Commercial Laboratories as Assessed by the Prospective Registry of Multiplex Testing. J Clin Oncol. 2016;34(34):4071–4078. [PubMed: 27621404]

16. Nussbaum RL, Yang S, Lincoln SE. Clinical Genetics Testing Laboratories Have a Remarkably Low Rate of Clinically Significant Discordance When Interpreting Variants in Hereditary Cancer Syndrome Genes. J Clin Oncol. 2017;35(11):1259–1261. [PubMed: 28135136]

17. Yang S, Cline M, Zhang C, Paten B, Lincoln SE. Data Sharing and Reproducible Clinical Genetic Testing: Successes and Challenges. Pac Symp Biocomput. 2016;22:166–176.