

Genetic Database Recognition Package Cover Sheet

Date: June 22nd 2018

Subject: **Genetic Variant Database Recognition Request**

Attention: **Brittany Schuck, Ph.D.**

Division of Chemistry and Toxicology Devices

Submission Type: **Q-SUBMISSION: INFORMATIONAL MEETING**

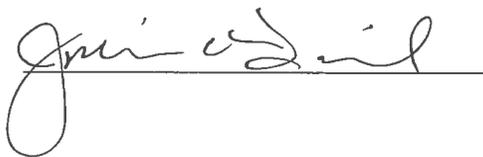
The Clinical Genome Resource consortia, ClinGen, requests recognition for the full dataset of germline variant classifications as outlined in the enclosed application. The ClinGen database variant classifications and the processes to support them encompass germline variants for hereditary diseases.

The point of contact for this submission is the ClinGen database coordinator.

Julianne M. O'Daniel
jodaniel@med.unc.edu
919-843-3702

Submitted on behalf of ClinGen.

The submitter, Julianne O'Daniel, believes, to the best of his or her knowledge, that all information submitted are truthful and accurate and that no material fact has been omitted.

A handwritten signature in black ink, appearing to read "Julianne O'Daniel", is written over a horizontal line.

ClinGen Human Variant Database

Contents

Background	2
Definitions	3
Supplemental files list	5
1.Database Procedures and Operations	6
2. Data Quality	12
3. Variant Evaluation and Assertions	16
4. Professional Training and Conflicts of Interest	24

The point of contact for this submission is the ClinGen database coordinator.

Julianne M. O'Daniel
jodaniel@med.unc.edu
919-843-3702

Submitted on behalf of ClinGen.

Background

Genomic variants are increasingly being associated with phenotypes and clinical conditions. This knowledge is already being used in clinical care, particularly to accurately diagnose inherited disorders. However, the aggregation of evidence and methods for assessing genes and variants and effective dissemination of this knowledge is needed to inform the clinical significance of the thousands of known genes and variants associated with disease to achieve greater integration into clinical care.

The Clinical Genome Resource (ClinGen) is a National Institutes of Health (NIH)-funded resource which aims to aggregate, curate, and make publicly available information pertaining to the clinical significance of genes and variants. Composed of 500+ contributors representing numerous aspects of human genomics and medical genetic stakeholder groups, ClinGen aims to standardize clinical annotation and interpretation of genes and genomic variation.

Key to the mission of ClinGen is transparent access to validated evidence. As such, the ClinGen website, accessed at: <https://clinicalgenome.org/>, contains access to all the ClinGen resources described herein.

This application pertains to the human variant interpretation activities of ClinGen. These activities are organized into Variant Curation Expert Panels (VCEPs) who carry out robust, variant assessment and classification in accordance with highest professional standards. The development, coordination and maintenance of VCEPs require step-wise approval in a controlled process under the purview of ClinGen oversight committees. Only variants that have been fully assessed by these ClinGen governed VCEPs are proposed for inclusion in the ClinGen Expert Curated Human Variant Database (“Database”) that we submit to the FDA for recognition. As of June 2018, ClinGen has 20 variant curation expert panels in the approved or developing stage.

Maintenance/Administration of the Database, and the processes outlined within this application will be overseen by the ClinGen Steering Committee and coordinated by a designated Database coordinator. The current Database coordinator is Julianne O’Daniel (jodaniel@med.unc.edu).

Definitions

ClinGen: The Clinical Genome Resource or “ClinGen” is a National Institutes of Health (NIH)-funded, multicenter consortium dedicated to building an authoritative central resource that defines the clinical relevance of genes and variants for use in precision medicine and research. The activities are directed through grants to the following centers: Baylor College of Medicine, Geisinger Health System, Harvard, Kaiser Permanente, Stanford, University of North Carolina, and the American College of Medical Genetics & Genomics (ACMG). Primary funding is through the National Human Genome Research Institute (NHGRI) with additional support for content curation from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD). <https://clinicalgenome.org/>

ClinVar: ClinVar is a funded resource of the National Center for Biotechnology Information (NCBI). The purpose for ClinVar is to provide a freely accessible, public, tracked record of reported relationships between human variation and observed health status with supporting evidence. ClinVar processes and publishes variant-level submissions in a standard format that includes the submitter, date of assessment, and accompanying evidence that supports the asserted pathogenicity classification. The alleles described in submissions are mapped to standard reference sequences and reported according to the Human Genome Variation Society (HGVS) standard. ClinVar data is accessible online for interactive users as well as those wishing to use ClinVar in daily workflows and other local applications. <https://www.ncbi.nlm.nih.gov/clinvar/>

ClinGen Expert Curated Human Variant Database (“Database”): The ClinGen Expert Curated Human Variant Database is referenced as the “Database” throughout this application. The Database refers to the collection of variants and the accompanying variant-specific assertion of pathogenicity classification, and supporting evidence summary, that have been approved as final from a ClinGen VCEP.

Steering Committee: The ClinGen Steering Committee is responsible for establishing standards, and oversight of all processes within the scope of ClinGen. Steering Committee membership is composed of the principal investigators at each of the funded ClinGen institutions as well as leadership from various ClinGen working groups and committees and the NHGRI program officer. Meetings are held monthly.

Clinical Domain Working Groups Oversight Committee (CDWG Oversight Committee): The ClinGen Clinical Domain Working Groups Oversight Committee (CDWG Oversight Committee) has responsibility for the general oversight of the development, approval and coordination of VCEPs. Membership in the CDWG Oversight Committee includes representatives from the Steering Committee, VCEP leaders and coordinators.

Variant Curation Expert Panel (VCEP): The ClinGen Variant Expert Panel (VCEP) references those groups who abide by the full ClinGen governing processes and thus follow a rigorous approval process. VCEPs are limited in scope to variant assessment for a single gene or set of related genes, defined by a disease phenotype. Membership of VCEPs is expected to represent the diversity of expertise in the specified gene/disease, including clinical, diagnostic laboratory, and research expertise. In addition to national and internationally recognized domain experts, VCEPs also include ClinGen trained biocurators to assist with the curation of variants to be assessed.

Variant Curation Interface (VCI): The Variant Curation Interface (VCI) is open source software developed by ClinGen. The source code can be accessed at <https://github.com/ClinGen/clincoded>. The VCI

supports the manual, expert curation of variant information through aggregation of external evidence about variants, organized based on standard evidence categories and classification criteria. The VCI further provides a means for evidence curated and pathogenicity classifications to be versioned and stored. <https://curation.clinicalgenome.org/>

ClinGen Allele Registry: The ClinGen Allele Registry (<https://reg.clinicalgenome.org>) provides unique identifiers (“CAids” or “canonical identifiers”) for genetic variants. These identifiers and additional web services provided by the Registry ensure unique naming and consistent identification of genetic variants. The Registry provides web services via User Interfaces and APIs that help retrieve an existing identifier or assign a new identifier based on unique combinations of variant attributes. The Registry also links to alternate identifiers already established in major public databases.

ClinGen Evidence Repository: The Evidence Repository provides programmatic access to machine-readable variant interpretations and supporting evidence in the ClinGen Expert Curated Human Variant Database (“Database”). The Repository is built on FAIR (“Findable”, “Accessible”, “Interoperable”, Reusable”) principles, allowing the community to reuse the machine-readable content either by downloading it completely and then ingesting into their own system or accessing it selectively via the Evidence Repository APIs. The Evidence Repository will be made accessible at <https://erepo.clinicalgenome.org>.

Sequence Variant Interpretation working group (SVI): The Sequence Variant Interpretation (SVI) Working Group is an internal ClinGen group led by experts in human variant interpretation. The purpose of the SVI is to support the continuous, evidence-based refinement and evolution of the published ACMG/AMP Interpreting Sequence Variant Guidelines to develop standard, quantitative approaches to variant interpretation. The SVI also provides training and consultative support to each VCEP for the development of consistent and harmonized gene/disease-specific refinements of the ACMG/AMP Interpreting Sequence Variant Guidelines.

ACMG/AMP Interpreting Sequence Variant Guidelines: The ACMG/AMP interpreting sequence variant guideline is the overarching guidance for the variant evaluation processes for Mendelian disease variants. This document is formally titled: “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.” It is available at: https://www.acmg.net/docs/standards_guidelines_for_the_interpretation_of_sequence_variants.pdf

This practice standard outlines the types of evidence that should be aggregated and assessed for robust interpretation of human variants and provides examples of sources to utilize for collection of the required evidence. It further provides criteria for combining evidence categories to classify the pathogenicity of variants for Mendelian disease according to five classes: pathogenic, likely pathogenic, uncertain significance, likely benign, and benign. The “ACMG/AMP criteria” refers to the criteria for the classification of variants based on the evidence categories met. The “refinement or specification of the ACMG criteria” refers to the refinement of evidence criteria to meet a particular evidence category as well as the refinement of the weight given to a particular evidence category in the classification of variant pathogenicity.

Supplemental files list

Cardiomyopathy VCEP Documentation

- Cardiomyopathy VCEP application
- Cardiomyopathy Variants
- Cardiomyopathy – MYH7 Criteria Summary
- Cardiomyopathy VCEP publication - Adaptation and validation of the ACMG-AMP criteria
 - Available online at: <https://www.nature.com/articles/gim2017218>

RASopathy VCEP Documentation

- RASopathy VCEP application
- RASopathy Variants
- RASopathy VCEP publication – ClinGen RASopathy Expert Panel consensus methods for variant interpretation
 - Available online at: <https://www.nature.com/articles/gim20183>

PAH VCEP Documentation

- PAH VCEP application
- PAH Variants
- PAH_ACMG_Specifications

VCEP Formation – Guidance SOP for formation, and the specification process for ACMG/AMP pathogenicity classification criteria. Available online at: <https://www.clinicalgenome.org/expert-groups/>

- Clingen_expert_panel_application
- Clingen_expert_panel_toolkit

ACMG-AMP Publication.Standards_guidelines...

- Protocol that governs all variant aggregation, curation and assessment processes
- https://www.acmg.net/docs/standards_guidelines_for_the_interpretation_of_sequence_variants.pdf

VCI Curation Help Document

- Also available online at: <https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help>

SVI Publications

- SVI publication_ACMG-AMP reputable source criteria
- SVI publication_Modeling ACMG-AMP classification as Bayes

ClinGen VCEP Conflict of Interest Survey

ClinGen Disclaimer Language

List of manuscripts in preparation, submitted and in press

1.Database Procedures and Operations

1.1 Transparency and Public Accessibility

Transparency and public access are a central component of ClinGen and as such all final assertions regarding variant pathogenicity, classification criteria, and curated evidence will be made available for any variant contained within the Database.

Variant-level Pathogenicity Assertions. All final assertions regarding the pathogenicity of a genomic variant will be made publicly available via submission to ClinVar. ClinVar is a freely accessible, National Center for Biotechnology Information (NCBI) maintained, public archive for externally submitted reports <https://www.ncbi.nlm.nih.gov/clinvar/variation/40702/> of human genomic variants and their relationship to human health. See <https://www.ncbi.nlm.nih.gov/clinvar/intro/> for reference.

The ClinVar variant record includes the following:

- Pathogenicity assertion using standard language: Benign, Likely Benign, Uncertain Significance, Likely Pathogenic, Pathogenic
 - Disease or condition name and inheritance pattern for which the pathogenicity assertion applies
 - Summary of the asserted pathogenicity classification and the evidence categories applied
 - Date of evaluation
 - Name of ClinGen Variant Curation Expert Panel that made the assertion
- See <https://www.ncbi.nlm.nih.gov/clinvar/variation/43088/#summary-evidence> for an example of a ClinVar variant page record from the Inherited Cardiomyopathy ClinGen Variant Curation Expert Panel.

The ClinVar submitter record will include the following:

- Name of the ClinGen Variant Curation Expert Panel (VCEP)
 - Contact information (name, phone number and email) for the VCEP
 - Hyperlink to the specified variant pathogenicity assertion criteria embedded within the ClinVar domain
 - Hyperlink to the VCEP initial membership and validation evidence for the specified variant pathogenicity assertion criteria embedded within the ClinVar domain
 - Hyperlink to the VCEP page on the ClinGen domain website
- See <https://www.ncbi.nlm.nih.gov/clinvar/submitters/506161/> for an example of a ClinVar submitter page record from the Inherited Cardiomyopathy ClinGen Variant Curation Expert Panel.

Assertion Criteria. The specified variant pathogenicity assertion criteria used for the interpretation of a variant will be publicly available via the ClinVar variant and VCEP submitter pages housed within the ClinVar domain as outlined above.

A more in-depth explanation of the development and validation of the assertion criteria is publicly available on the specific VCEP page within the ClinGen domain. Additional supporting evidence includes:

- Listing of VCEP leaders and coordinators including contact information
- Full listing of current membership and their employment affiliations
- Outline of assertion criteria development and validation
- Hyperlink to resulting peer-reviewed publications of the VCEP's assertion criteria

See <https://www.clinicalgenome.org/working-groups/clinical-domain/cardiovascular-clinical-domain-working-group/myh7-variant-curation-expert-panel/> for an example of a VCEP page housed within the ClinGen domain.

Curated Evidence. The evidence curated and assessed to make a pathogenicity assertion is stored within the ClinGen Variant Curation Interface (VCI). The VCI software is open source software (<https://github.com/ClinGen/clincoded>) developed by ClinGen. The VCI is password controlled and can be accessed at: <https://curation.clinicalgenome.org/>. Access to versioned, curated evidence in the VCI is available on request directly through the VCEP group leader and/or coordinator or through individual password request.

In addition, a summary of the evidence evaluated, and criteria applied to classify a variant will accompany all final pathogenicity assertions for Database variants when submitted to ClinVar. The summary includes at a minimum: a list of the specific variant interpretation criteria applied based on the underlying disease-specified ACMG/AMP clinical interpretation guidelines as well as the source of the evidence to apply those criteria. See example below. The content of this evidence summary is anticipated to change over time to reflect efforts to increase the amount of information described as well as respond to the preferences of external users such as physicians and genetic counselors.

This variant was classified using modified ACMG criteria (ClinGen Expert Panel; manuscript in preparation). Please see a summary of the criteria and criteria codes in the "ACMG variant classification (RASopathy)" document (assertion method column). The following criteria were met: PM1, PM2, PM6_Strong, PP2, PP3, PS4_Supporting. Additional case-level data provided by: SCV000203922, SCV000207755, Cave et al.

Programmatic access to machine-readable content. Machine-readable variant interpretations and supporting evidence produced via the VCI and stored in the Database will also be made accessible via the Evidence Repository (to be made accessible at <https://erepo.clinicalgenome.org>). The Repository is built on FAIR ("Findable", "Accessible", "Interoperable", Reusable") principles, allowing the community to reuse the machine-readable content either by downloading it completely and then ingesting into their own system or accessing it selectively via the Evidence Repository APIs.

1.2 SOP Version Control

There are several processes that guide different aspects of how variant information is aggregated, curated and evaluated. All SOPs that govern these processes will be made publicly available in their current version. Prior versions will be made available and can be requested via the relevant ClinGen contact. Changes will be clearly documented through highlight, footnote, or other formatting tool within the new version.

Current Standard Guidance documents and locations:

- VCEP Formation – Guidance SOP for formation, and the specification process for ACMG/AMP pathogenicity classification criteria
 - <https://www.clinicalgenome.org/expert-groups/>
- ACMG/AMP Variant Interpretation Guideline – Protocol that governs all variant aggregation, curation and assessment processes
 - https://www.acmg.net/docs/standards_guidelines_for_the_interpretation_of_sequence_variants.pdf
- VCEP-specified ACMG/AMP Criteria – The specified protocols are available in three locations:
 - (1) Accompanying variant assertions in ClinVar
 - (2) The VCEP Submitter page within ClinVar
https://www.ncbi.nlm.nih.gov/clinvar/docs/submitter_list/
 - (3) The VCEP Work Group page within ClinGen
 - Variant Curation Interface – Document to guide VCI use
 - <https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help>
 - Variant Curation Expert Panel (VCEP) Methods – SOP for on-going variant curation. These will be posted as developed by each VCEP that has completed the 4-stage process for internal recognition.
 - <https://www.clinicalgenome.org/working-groups/clinical-domain/>

1.3 Data Preservation

All data connected to variants within the Database, including the variant-specific assertion of pathogenicity, supporting evidence summary, assertion criteria used, ClinGen Allele Registry identifier, and VCEP group responsible will be publicly available via ClinVar. As a NCBI maintained archival database for human variants, ClinVar, represents a stable public repository designed for data preservation. The ClinGen Evidence Repository will provide programmatic access to machine-readable variant interpretations and supporting evidence. The content of the Evidence Repository will be exportable in full as machine-readable text files in JSON-LD format. Designated releases will be exported for archival purposes and will be accessible at <https://erepo.clinicalgenome.org> and will be hosted at multiple commercial provider sites for redundancy.

All data contained within the ClinVar repository database is backed-up nightly and kept in short term storage for 60 days. Long term data is archived to tape, which will be kept indefinitely.

The VCI is backed up nightly on Amazon Web Services. Use of unique variant IDs and internal checks are performed to ensure accurate data are stored in the back-up files.

1.4 Security and Privacy

No personal/protected health information is connected or maintained within the VCI or attached to variants contained within the Database. Therefore, no protected health information will be part of the variant assertion. Documentation of HIPAA training as required by each VCEP member's home institution will be maintained by the relevant VCEP coordinator.

All coordinators and members of each VCEP receive training on the use of the VCI. This training will include the policy that only data stripped of any possible protected health information will be entered in the VCI. Any and all evidence curated and reviewed as part of the assessment process should be collected from current published sources (e.g. PubMed scientific articles or online public research databases such as Leiden Online Variant Database) or from de-identified clinical lab databases which are required to meet federal HIPAA laws.

Variant evidence aggregated from external sources or curated within the VCI is reviewed by the individual VCEP for the purpose of evaluation and determination of final pathogenicity classification. All curation activities within the VCI are tracked based on the user's login credentials. This allows audit of VCI users at any time. If protected health information is observed during this variant information review: (1) the protected health information will be removed by the coordinator or appropriate VCEP leader; and (2) the VCEP members will be re-trained on the ClinGen policy not to store protected health information.

1.5 Data Formats and Nomenclature

Standard nomenclature will be used across all processes of the variant curation, assessment and publication of final assertion. As the standards for representing variant knowledge in machine-readable formats are under development, we will continue to adopt the latest standards developed by the community and will also actively participate in shaping these standards in order to facilitate exchange, archiving, and dissemination of variant knowledge across the user community.

Within ClinVar, all data submissions must meet the following standard nomenclature formats: Human Genome Variation Society (HGVS) nomenclature for variants including a reference transcript (RefSeq ID) and/or genomic coordinates according to GRCh37 or GRCh38 reference assemblies, HUGO gene name (HGNC gene symbol), associated condition names using MedGen terms (ClinGen uses MONDO IDs which are then mapped to MedGen terms), standard ACMG/AMP recommended classification terms. MONDO, or the Monarch Disease Ontology, identifiers are available at <https://www.ebi.ac.uk/ols/ontologies/mondo>.

ClinGen is defining and representing variants in a normalized manner in order to achieve interoperability and thus facilitate aggregation and exchange of variant knowledge. The first product of the ClinGen is the Variant Model that has been first implemented in the ClinGen Allele Registry and will be used to represent variants in the VCI and the Evidence Repository. Members of the ClinGen project have now expanded this standardization by involving the Global Alliance for Genomics and Health (GA4GH) Variation Modeling Stream. We anticipate modifying the original ClinGen Variant Model to achieve adoption by the GA4GH.

2. Data Quality

All variants within the Database and data supporting the genotype-phenotype assertions of pathogenicity within the Database represent application of the best community practice with regard to the aggregation, curation and assessment of human genomic variants and their relationship to diseases and traits. To ensure this high quality, ClinGen utilizes an organization-wide system of method and disease expertise.

The Sequence Variant Interpretation (SVI) workgroup consists of experts in the methods of human variant interpretation, and in particular the application of the ACMG/AMP criteria, the current accepted standard for this activity. As referenced throughout, variant interpretation activities follow the ACMG/AMP human variant interpretation guidance which represents the joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. The SVI is responsible for providing guidance to ClinGen for all variant assessment activities. This includes education for the VCEPs regarding the ACMG/AMP criteria, highlighting where gene/disease-specific clarifications may be appropriate.

The Variant Curation Expert Panels (VCEPs) are composed of scientific expertise regarding gene function and clinical expertise regarding disease manifestations. Further, each VCEP includes biocurators, who have been trained in the use of the VCI and can facilitate the curation of variant-level information.

The Variant Curation Interface (VCI) aggregates external evidence about variants and supports the manual curation of variant information. Supporting evidence is accessible to all VCI users. The data fields within the VCI are defined by the evidence categories provided by the ACMG/AMP guideline. Dynamic links to external information sources are also embedded within the relevant evidence tabs. External information sources are defined for each evidence category as outlined in the ACMG/AMP guidance document as well as additional resources identified as valid sources by the SVI.

Evidence categories and example sources are below. Population, in-silico, and gene-centric evidence is aggregated and displayed for users. Experimental and case-level evidence is manually curated and entered by the user from published literature and other sources by the curator.

- Population-based data (e.g. allele frequencies in various populations)
 - gnomAD, ExAC, PAGE, 1000 Genomes, Exome Sequencing Project and curated literature
- In silico prediction model data (e.g. evolutionary conservation, splicing predictors)
 - REVEL, SIFT, PolyPhen2, LRT, FATHMN, CADD, phyloP100way, GERP++, MaxEntScan, NNSPLICE
- Experimental data (e.g. animal model, tissue expression)
 - Curated literature for functional domain and experimental evidence
- Case-level data
 - Curated literature for case-control, healthy population observations
 - Curated literature for segregation, cis-/trans-, de novo, and phenotypic specificity

- Curated non-published clinical data from diseased individuals (e.g. from laboratory testing)
- Gene-centric clinical validity data
 - ExAC constraint scores
 - Links to gene resources such as HGNC, Entrez Gene, Ensembl, UniProtKB and ClinVar

Standard machine-readable formats for provenance and evidence metadata are key to reusability of ClinGen knowledge. ClinGen's growing collaboration with the Monarch Initiative for both the ontological representations of disease/phenotype models as well as the emerging Scientific Evidence Provenance Information Ontology (SEPIO) have enabled ClinGen DMWG to release the first draft of the ClinGen SEPIO variant interpretation model. This model will be implemented in the VCI and Evidence Repository, to enable dissemination of variant knowledge in standard machine-readable format. ClinGen will continue to engage in standards development with SEPIO, GA4GH, and other community-wide efforts to define and implement standards for the exchange, archiving, and dissemination of variant knowledge in machine-readable formats.

2.2 Data Uniqueness

A variant assertion requires an assessment of both (1) whether a variant disrupts the function of a gene and (2) whether that disruption causes or contributes to a particular disease or trait. Therefore, a specific sequence variant may have more than one assertion based on the assessment of more than one possible disease or trait. Unique variant-identifiers are supported through several measures: (1) ClinGen Allele Registry identifier, (2) Variant Curation Interface record, and (3) ClinVar submission accession number (SCV).

ClinGen Allele Registry. First, each specific human genomic sequence variant is required to have a unique variant identifier. For some variants a ClinVar Variation ID number exists and will suffice. In addition, every variant will be assigned a canonical identifier by the ClinGen Allele Registry. If a variant has both identifiers, the two will be in one-to-one correspondence and they will be cross-referenced by the two sources. Canonical identifiers are based on sequence alignment against genomic and transcript references. The Registry provides unique variant identifiers both programmatically (via APIs) and via a public interface accessible at <http://reg.clinicalgenome.org>.

The current content of the Registry is searchable using Human Genome Variant Society (HGVS) expressions representing nucleic acid or protein changes across more than 500,000 reference sequences (genome assemblies, transcripts, amino acid sequences). Alleles can be also queried by locus, gene or ClinVar or dbSNP identifiers. The registry regularly imports variants from ExAC, ClinVar and other databases. To generate Canonical identifiers, the Registry maps variants algorithmically across known reference sequences.

If a variant is not present in the Registry, authorized users may register the variant and receive an assigned identifier within seconds. To register new alleles in the Allele Registry, users need a valid login and password which can be obtained by an email request to bcm.clingen@gmail.com.

Variant Curation Interface (VCI) Record. The VCI creates a unique variant interpretation record based on the Allele Registry identifier and, as appropriate, the disease/condition name as selected from standard MONDO terminology.

Once a variant assertion has been submitted to ClinVar, the resulting ClinVar accession number (ClinVar Submission Record or SCV number) is then linked to the VCI record through the unique identifier for the approved interpretation. This link assures that any future revisions to an assertion will be linked to the public assertion in ClinVar.

ClinVar Submission Record (SCV). Upon successful submission to ClinVar, each variant-level record will be assigned an accession number of the format SCV000000000.0. If there are multiple submissions for the same variant/condition relationship, they are aggregated within ClinVar's data flow and reported as a reference accession of the format RCV000000000.0. Because of this model, one variant will be

included in multiple RCV accessions whenever different condition names are reported for the same variant allele.

Further, if a VCEP includes data from a separate ClinVar submitter to support a variant assertion, the SCV number is noted to prevent future redundant use of the same evidence. This notation occurs within the supporting evidence summary that accompanies the variant assertion and will be included in the VCI within the case-level data for that variant record.

3. Variant Evaluation and Assertions

The overarching guidance for the variant evaluation processes for Mendelian disease variants is the ACMG/AMP guideline for variant interpretation. This document, published in 2015, is formally titled: “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.” It is available at:

https://www.acmg.net/docs/standards_guidelines_for_the_interpretation_of_sequence_variants.pdf

This practice standard outlines the types of evidence that should be aggregated and assessed for classifying the pathogenicity of human variants for Mendelian disease according to five classes: pathogenic, likely pathogenic, uncertain significance, likely benign, and benign. It further provides examples of sources to utilize to collect the required evidence.

3.1 Evaluation and Data Versioning within the Variant Curation Interface (VCI)

As described previously (section 2.1 Metadata), the aggregation and curation of variant information is performed within the VCI. Evidence categories are established based on those set forth by the ACMG/AMP guideline. The structure of the VCI is defined by those evidence categories. Population, in silico and gene-centric evidence is aggregated and displayed for the user. When evidence is unavailable for display via API, dynamic links to external information sources are embedded within the relevant evidence tab. External information sources are defined based on current practice standards as communicated by the ClinGen SVI. The VCI will programmatically incorporate any future amendments and updates made to the ACMG/AMP variant interpretation guideline or additional modifications and specifications made by ClinGen’s Sequence Variant Interpretation guidelines.

A calculator programmed within the VCI explicitly applies the criteria for combining criteria evidence codes to classify variants as outlined in Table 5 of the referenced ACMG/AMP guidelines (see below). However, the calculated pathogenicity does not equate to a final assertion of pathogenicity and curators are able to modify the calculated pathogenicity, providing a reason for its modification

The VCEP responsible for that particular variant assesses the full spectrum of evidence in the context of the particular gene/disease and their specified ACMG/AMP criteria. The resulting pathogenicity assertion is marked as “Approved” within the VCI such that the “Approved” assertion represents the current assertion in the VCI. All data contained within the VCI variant record at that time is captured and stored as a digital snapshot. The variant record is then viewable to all users within the VCI. New evidence may be added for the variant in the VCI at any time. If an updated variant evaluation (aka “reevaluation”) is desired, the formal curation, evaluation and pathogenicity classification process must proceed afresh through the standard approval process. All current and previously approved classifications are viewable in the VCI as classifications are versioned rather than overwritten.

Table 5 Rules for combining criteria to classify sequence variants

Pathogenic	<ul style="list-style-type: none"> (i) 1 Very strong (PVS1) <i>AND</i> <li style="padding-left: 20px;">(a) ≥ 1 Strong (PS1–PS4) <i>OR</i> <li style="padding-left: 20px;">(b) ≥ 2 Moderate (PM1–PM6) <i>OR</i> <li style="padding-left: 20px;">(c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) <i>OR</i> <li style="padding-left: 20px;">(d) ≥ 2 Supporting (PP1–PP5) (ii) ≥ 2 Strong (PS1–PS4) <i>OR</i> (iii) 1 Strong (PS1–PS4) <i>AND</i> <li style="padding-left: 20px;">(a) ≥ 3 Moderate (PM1–PM6) <i>OR</i> <li style="padding-left: 20px;">(b) 2 Moderate (PM1–PM6) <i>AND</i> ≥ 2 Supporting (PP1–PP5) <i>OR</i> <li style="padding-left: 20px;">(c) 1 Moderate (PM1–PM6) <i>AND</i> ≥ 4 supporting (PP1–PP5)
Likely pathogenic	<ul style="list-style-type: none"> (i) 1 Very strong (PVS1) <i>AND</i> 1 moderate (PM1–PM6) <i>OR</i> (ii) 1 Strong (PS1–PS4) <i>AND</i> 1–2 moderate (PM1–PM6) <i>OR</i> (iii) 1 Strong (PS1–PS4) <i>AND</i> ≥ 2 supporting (PP1–PP5) <i>OR</i> (iv) ≥ 3 Moderate (PM1–PM6) <i>OR</i> (v) 2 Moderate (PM1–PM6) <i>AND</i> ≥ 2 supporting (PP1–PP5) <i>OR</i> (vi) 1 Moderate (PM1–PM6) <i>AND</i> ≥ 4 supporting (PP1–PP5)
Benign	<ul style="list-style-type: none"> (i) 1 Stand-alone (BA1) <i>OR</i> (ii) ≥ 2 Strong (BS1–BS4)
Likely benign	<ul style="list-style-type: none"> (i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) <i>OR</i> (ii) ≥ 2 Supporting (BP1–BP7)
Uncertain significance	<ul style="list-style-type: none"> (i) Other criteria shown above are not met <i>OR</i> (ii) the criteria for benign and pathogenic are contradictory

Table from Richards 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* 2015

3.2 Development and Validation of Variant Evaluation Protocols (aka ACMG/AMP criteria specification)

All curation and variant interpretation will be carried out by members within ClinGen approved Variant Curation Expert Panels (VCEPs). The VCEP development and criteria validation process is outlined in the following guidance documents:

- VCEP Development – Guidance SOP for forming a VCEP and specifying the ACMG/AMP criteria for individual genes and diseases
 - <https://www.clinicalgenome.org/expert-groups/>

These documents, referred to as the Expert Panel Tool Kit serve as the SOPs for the development of all ClinGen-formed VCEPs. Current versions of these documents are included in the Supplemental Material. These guiding documents are updated through the Clinical Domain Working Group Oversight Committee with overarching guidance from the ClinGen Steering Committee. Current versions are publicly available at the link above for review. Past versions are made available on request.

Each VCEP is tasked with the process of development and validation of specifications to the ACMG/AMP variant interpretation criteria based on the unique characteristics of their particular disease and/or genes(s). An overview of the entire process is depicted in the figure on page 20. Further a peer-reviewed manuscript describing the VCEP approval process has been accepted for publication in *Human Mutation*.

Step 1- VCEP Formation. The first step for each VCEP is the identification of membership and scope, with input from their parent Clinical Domain Working Group (CDWG), if applicable. For VCEPs formed outside of the current CDWGs, the Clinical Domain Work Group (CDWG) Oversight Committee may also provide guidance as needed to ensure appropriate expertise is represented. Membership affiliations are collected, and all members must complete Conflict of Interest (COI) documentation. This is described in Section 4.2, Conflict of Interest. The VCEP submits a Step 1 application for approval by the CDWG Oversight Committee before proceeding to Step 2.

Step 2- Develop Draft Variant Classification Criteria. The SVI work group provides education and training for VCEP members as described in section 4.1 to ensure VCEPs understand the ACMG/AMP variant interpretation guidance, classification criteria, and evidence categories. Building on this foundation, VCEPs receive training in the criteria specification process, focusing on criteria and evidence types that should be considered for disease/gene(s) specification and those that are broadly applicable across domains, such as applying statistical rigor to calculate allele frequency thresholds for benign classifications. The VCEP develops an organizational strategy to discuss, determine, and draft specifications to the assertion criteria using the ACMG/AMP criteria as their framework and utilizing the SVI for feedback and recommendations, as needed. A draft version of the specified ACMG/AMP criteria is submitted to the SVI for review and recommendations prior to pilot testing and validation.

Step 3 – Validation of Specified Criteria. Each VCEP then applies their specified ACMG/AMP criteria to a set of known variants. This example set could vary but should include variants that meet the following conditions:

- 1) At least 10 - 12 Pathogenic/Likely Pathogenic variants, 10 - 12 Benign/Likely Benign variants and 10 - 12 Uncertain Significance variants or those with conflicting interpretations in ClinVar; and
- 2) Include variants with a variety of different evidence types for the gene to test all relevant criteria codes.

The VCEP refines the criteria specifications based on test curations. If any significant refinements are made, the revised ACMG/AMP criteria specifications are resubmitted to the SVI working group.

Prior to final approval as a ClinGen VCEP, the developing VCEP must obtain SVI approval of their final criteria specifications and present a completed application, including plans for ongoing variant curation, review, and reanalysis, as well as discrepancy resolution to the CDWG Oversight Committee for approval. The application document is contained within the Expert Panel ToolKit at <https://www.clinicalgenome.org/expert-groups/>.

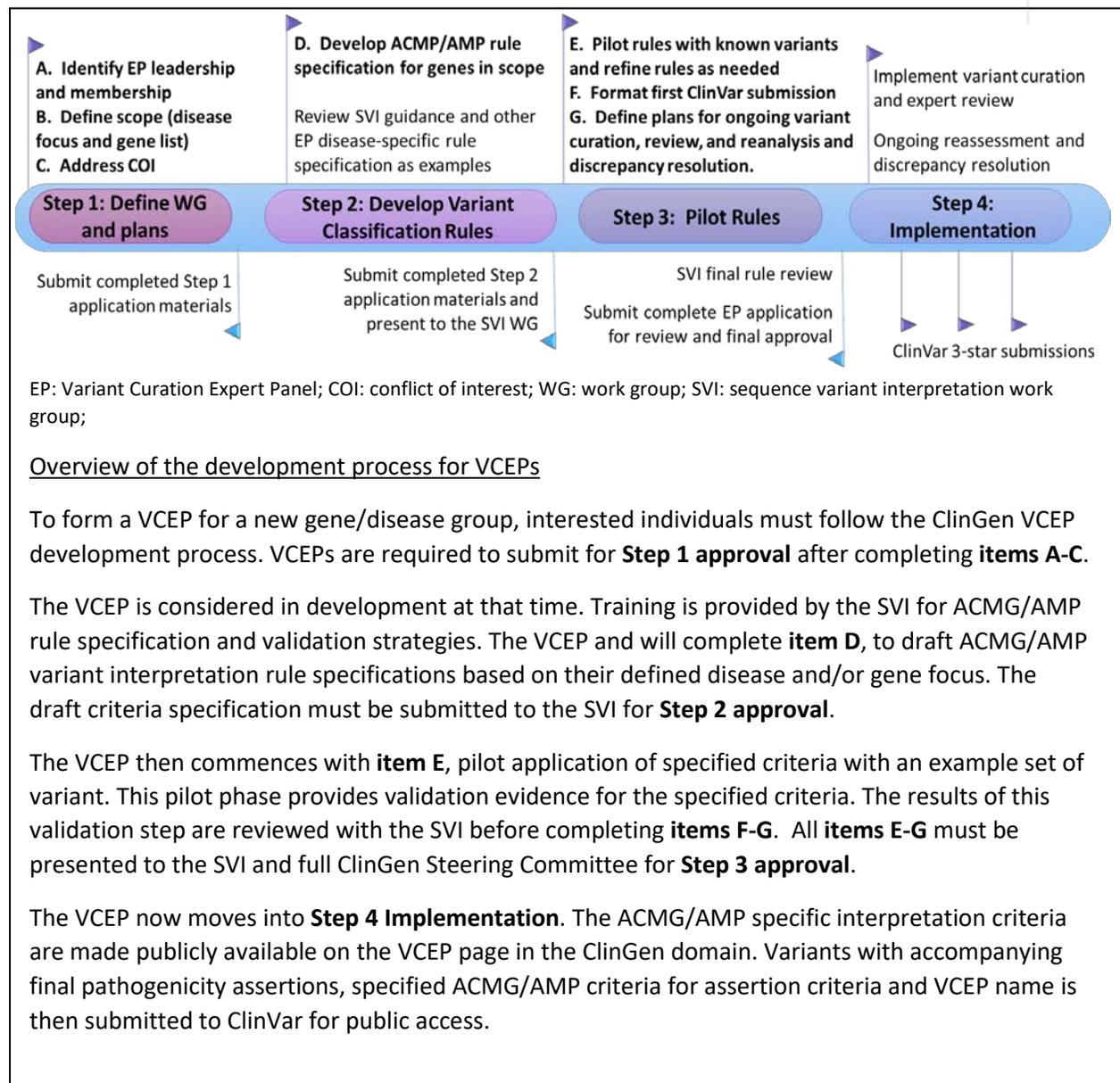
The completed application must contain:

- Membership listing, including affiliations and disclosure of COI
- Designated leadership and coordinator, including contact information
- Scope of the VCEP, including the disease and/or gene(s) covered
- Detailed description of the specified ACMG/AMP criteria, including
 - Pathogenicity assertion criteria, including description and explanation for any combinations of criteria and evidence sources that could be used to classify broad categories of variants (e.g. Benign or Likely Benign) in a batch.
 - Evidence and rationale to support the rule specifications. This explanation should include a plain language description of how the rule specification impacts the overall assertions (eg. Provides a more conservative population frequency such that fewer variants will meet the automatic criteria for Benign and thus require further evidence be curated and reviewed before a final assertion can be made).
 - Description and explanation for any combinations of criteria and evidence sources that could be used to classify broad categories of variants (e.g. Benign or Likely Benign) in a batch.
 - Summary of validation process, including a list of the example variant set and asserted pathogenicity*
- Standard Operating Procedures for variant curation and review
- Procedures for periodic reanalysis and discrepancy resolution

**The full validation evidence generated through the pilot phase will be made available by request to the VCEP leadership or coordinator.*

Once approved by the SVI and CDWG Oversight Committee, the full application document is made publicly available on the ClinVar Submitter page under “Expert panel documentation. See

<https://www.ncbi.nlm.nih.gov/clinvar/submitters/506439/> as an example. The now approved ClinGen VCEP is encouraged to publish their gene/disease-specific ACMG/AMP criteria specification and validation in a peer-reviewed journal.



3.3 Application of Specified ACMG/AMP Criteria for Variant Evaluation

Step 4 – Implementation. Following approval of the VCEP full application, the VCEP then proceeds with application of the specified ACMG/AMP assertion rule criteria for variant interpretation within their defined disease/gene scope. All variants curated, evaluated and assigned a final pathogenicity assertion during the validation process are submitted to ClinVar. This ClinVar submission follows the process as outlined in section 1.1 Transparency.

The ongoing variant curation and evaluation process is directed within each VCEP following the general ClinGen VCEP Variant Curation Guidance (below).

General ClinGen VCEP Variant Curation Guidance

Selection of Variants

Selection of variants for curation and evaluation will be made by consensus discussion by the VCEP membership. Variants are prioritized based on frequent reports in current clinical databases such as ClinVar as well as those nominated by external parties.

Variant Curation and Preliminary Evaluation

Each VCEP will define which method works best for proceeding with curation in their group. All approaches contain at least two reviewers. Acceptable methods include:

- One curator per variant reviews evidence and applies criteria to determine a preliminary assertion. This is followed by either (1) blinded expert review to validate the assertion, or through (2) full VCEP review.
- Each variant is assessed by 2 trained, independent curators to review evidence and apply criteria to determine a preliminary assertion. Discordant interpretations are then reviewed by the full committee.

The primary curator(s) will enter and review variant information for each evidence category in the VCI. The curator will utilize the VCEP-specified ACMG/AMP criteria to determine which of the ACMG/AMP rule codes to apply.

The VCI programmatically aggregates all of the applied codes and calculates a tentative variant classification of pathogenicity based on Table 5 from the ACMG/AMP variant interpretation paper (included on page 17 of this application). The curator will either select the calculated classification or select a different pathogenicity classification to assert. The curator will provide sufficient statements to support the selected classification, including a rationale if not applying the calculated classification.

Final Variant Review and Assertion of Pathogenicity Classification (Approval)

Final approval by the VCEP is required for all variant classifications. The approval of the classification is recorded in the VCI and the variant record is labeled as “Approved.”

ClinVar Submission

Upon finalization, the VCEP is responsible for coordinating the ClinVar submission process. The designated member will contact liaisons within the VCI and Data Model workgroups to ensure proper formatting of the variant data for submission. The VCEP designate provides final review of the formatted variant data, ensuring the final variant assertions as recorded in the VCI variant record are what will be submitted to ClinVar. Once approved for accuracy against the VCI record of final pathogenicity classifications, the variant data file is submitted to ClinVar.

VCEPs are expected to maintain ongoing curation and variant interpretation resulting in submissions to ClinVar at least once per year.

Distribution of machine-readable content via the ClinGen Evidence Repository

Variant interpretations that result from variant curation will also be made accessible programmatically in compliance with FAIR principles via the ClinGen Evidence Repository.

Review of specified ACMG/AMP Criteria

The VCEP will undergo a review of their specified ACMG/AMP criteria on an annual basis or as appropriate based on new gene-specific knowledge or SVI work group criteria guidance. This process will include consideration of any new guidance put forth by the SVI as well as updates to membership and scope, and in scientific and clinical knowledge about the characteristics of the particular disease/gene(s) group. Any changes/revisions made to the VCEP’s specified criteria must be approved by review of the SVI work group and the CDWG Oversight Committee.

3.4 Re-Evaluation of Variants with Final (Approved) Assertions

VCEPs are expected to keep their variant interpretations up-to-date and to expedite the reassessment of variants that have a conflicting assertion. All variant assertions are publicly accessible via ClinVar. This public access provides a means for encouraging inquiries and public comments regarding the evidence and rationale for classifications. To support the receipt of comments and inquiries, the contact information for each specific VCEP is made public on the ClinGen VCEP page as well as the ClinVar Submitter page. Users can quickly navigate to this contact information within ClinVar following imbedded links from the variant assertion page. Further, a mechanism to receive public feedback is available via a comment box present on every page of the ClinGen website.

Re-Evaluation

VCEPs are expected to keep their variant interpretations up-to-date and to expedite the reassessment of variants that have a conflicting assertion submitted to ClinVar. For resolution of conflicting classifications, VCEPs are expected to contact the submitter. Re-evaluation may also be initiated based on public inquiry and/or receipt of new data. Reassessment then proceeds as follows:

- Address newly submitted discrepant variant classifications, defined as those submitted to ClinVar after or up to 12 months prior to the initial VCEP submission
 - Discrepancy between Pathogenic/Likely Pathogenic and any other classifications should occur within 6 months
 - Discrepancy between Uncertain Significance vs Likely Benign/Benign should occur within 2 years
- Address all public inquiries and/or new or alternative evidence submitted to the VCEP
- Reassess all VCEP variants classified as Likely Pathogenic or Uncertain Significance at least every 2 years
- Reassess all VCEP variants classified as Likely Benign whenever new large population datasets are released

4. Professional Training and Conflicts of Interest

4.1 Professional Training

ClinGen VCEPs are expected to represent the diversity of expertise in the field, including all major areas of expertise (clinical, diagnostic laboratory, and research). Membership should include representation from three or more institutions and will encompass disease/gene expert members as well as biocurators. Biocurators do not have to be gene/disease experts and will be primarily responsible for assembling the available evidence for subsequent expert member review.

Biocurator Proficiency Training

Biocurators will have variable levels of variant interpretation experience. Each VCEP is responsible for coordinating and monitoring training and proficiency of their biocurators in procuring the appropriate data, assessing the data in the context of variant interpretation, and entering the data with sufficient detail into the VCI. New biocurators and/or those biocurators deemed by the VCEP to require additional training are paired with an experienced VCEP biocurator who can teach the relevant skills as they go through the variant interpretation process together.

To facilitate the training process:

1. ClinGen will provide general training materials for variant interpretation to new biocurators and/or newly forming VCEPs.
2. Each VCEP may provide additional training specific to the gene(s) they are working on.
3. Biocurators are expected to join the ClinGen Biocurator working group. This group provides a forum for training and education of all ClinGen gene and variant biocurators. The group meets bimonthly by conference call. Topics on the calls range from SVI updates, use of the ClinGen interfaces (including VCI), and presentations from experts on relevant resources such as Ensembl, ClinVar, gnomAD etc. Recordings of all calls and relevant slide sets are archived on the Biocurator Working Group page on ClinGen's Confluence site so that they can be accessed at any time. The listserv for the Biocurator Working Group provides a mechanism for disseminating updates, such as new guidance from the SVI, to ensure that biocurators stay current with ClinGen best practices.
4. A help document on how to use the Variant Curation Interface is available at <https://github.com/ClinGen/clin coded/wiki/VCI-Curation-Help>. Biocurators new to the VCI can gain experience by doing practice curations in the VCI as a trainee delegate.

After this training period, the proficiency of the biocurator is evaluated by the VCEP. For example, a biocurator might be asked to independently collect the relevant data and interpret a variant in the VCI. VCEPs may develop a training set of variants for this purpose or may compare the new biocurator's assessment with that of an experienced VCEP member. Once a biocurator has achieved proficiency, this is documented by the VCEP coordinator. As previously noted, all variant interpretations are reviewed by the full VCEP membership including non-biocurator, clinical or disease experts prior to being approved.

HIPAA Training

All members are responsible for obtaining HIPAA and human subjects training based on their home institutional/affiliation guidelines and the level of access to human subject data. VCEP leadership and coordinators ensure that members do not inadvertently share data that has not been stripped of protected health information or other identifiers. Identifiable and/or protected health information is not necessary for variant curation and evaluation and is not stored in the VCI or variant evidence summary.

Sequence Variant Interpretation Education and Training

The SVI work group provides ClinGen-wide guidance for the process of human variant evaluation to facilitate harmonization of approaches across ClinGen work groups and in particular the VCEPs. This guidance is formally presented at regular workgroup meetings via conference calls and published in peer reviewed literature.

SVI educational documents are publicly available on their ClinGen page at:

<https://www.clinicalgenome.org/working-groups/sequence-variant-interpretation/> Topically focused presentations are available on request.

All VCEPs receive VCEP-specific training and education along the following topics:

- ACMG/AMP criteria for variant interpretation
- ACMG/AMP criteria specification process, including education about which criteria are appropriate for a disease/gene-specific optimization process (eg. establishment of allele frequency cut-offs)
- Variant curation and use of the VCI

VCI representatives provide one-on-one training regarding the use of the VCI for variant curation.

Additional VCI training documents are publicly available for reference:

- Guidance for using the VCI is available through a “Learn More” link on the interface home page (<https://curation.clinicalgenome.org/>).
- Release descriptions of changes to the VCI software are versioned and can be found through the “Learn More” links at this url <https://github.com/ClinGen/clincoded/wiki/Releases>
- Curation activities within the VCI are guided by the VCI Curation Help documents. These are available through the “Learn More” pages at this url <https://github.com/ClinGen/clincoded/wiki/VCI-Curation-Help>

4.2 Conflicts of Interest

ClinGen VCEPs represent the diversity of expertise in the field, including clinical, diagnostic laboratory, and basic research. As such the VCEP is composed of a sufficient number of eligible expert reviewers to address academic and financial conflicts of interest that may arise. This number is typically ≥ 3 . Membership includes representation from three or more institutions and encompasses disease/gene expert members as well as biocurators. Biocurators do not have to be gene/disease experts and are primarily responsible for assembling the available evidence for subsequent expert member review. Biocurators will have at least a bachelor's level degree (though many have MS and PhD level degrees) and must complete proficiency training as outlined in section 4.1.

Expert Panels and their overarching Clinical Domain Working Groups are required to identify any conflicts of interest to ensure that members with academic or financial conflicts do not serve as the sole arbiter of gene or variant classifications for which they may have a biased perspective (e.g. if an individual published the first paper to implicate a gene in a disease). Each VCEP member is required to complete and submit a COI information form which is reviewed by the Clinical Domain Oversight Committee and maintained by the VCEP coordinator. This form is included in the Supplemental materials.

Possible COI:

- Academic COI: Authors of literature about relevant variants may serve on the VCEP and are welcome to voice their opinion but should not be the major arbiter of a variant classification when there is limited data available and it was provided by that individual or the individual's lab group.
- Financial COI: Commercial entities may participate on the VCEP but should not be the major arbiter of a variant classification when there is limited data available and it was provided by that entity.

Actions to address possible COI:

No special measures are needed if there is group consensus on a variant classification; however, if a vote is needed, those with relevant conflicts of interest recuse themselves.

All conflicts are declared publicly on the relevant VCEP group page at the clinicalgenome.org website and reported in publications as appropriate.