Implementation of Gene-specific ClinGen Variant Classification Recommendations using Artificial Intelligence: Frequency Thresholds

Yaron Einhorn, MSc;[1] Adaia Kamshov, MD, MSc;[1] Oron Lev, MSc;[1] Moshe Einhorn, MSc;[1] Nurit Paz-Yaacov, PhD;[1] Kamie Shami, BSc;[1]
Elaine Lyon, PhD, FACMG;[1,2] Susan Gross, MD, FACMG[1,3]

[1]Genoox, Tel Aviv, Israel; [2]Clinical Services Laboratory, HudsonAlpha Institute of Biotechnology, Huntsville, AL, USA; [3]Icahn School of Medicine at Mount Sinai, New York, NY, USA

ABSTRACT

Background: The 2015 ACMG/AMP sequence variant interpretation guidelines provided a framework for classifying variants based on several benign and pathogenic evidence criteria. Recently, ClinGen has introduced several expert gene/disease panels to establish more specific classification recommendations.

Aim: We hypothesized that artificial intelligence (AI)-based algorithms that are focused on implementation of gene-specific recommendations may aid in the process of variant curation.

Methods: The AI-based variant classification engine (aiVCE) was developed and trained using the curation processes recommended by the different expert panel (EP) teams for defining frequency-related rules (PM2, BS1, BS2, BA1). Pathogenic variants were extracted from ClinVar, and their corresponding population frequencies were gleaned from gnomAD. The aiVCE was benchmarked against the recently published gene-specific datasets employed to develop the recommendations by each EP, including MYH7, PTEN, CDH1, PAH, Rasopathy, and hearing loss genes, both in overall classification and in frequency rule thresholds.

Quality control experiments excluded variants with limited certainty for pathogenicity. Consistent with the EP groups, the threshold for BS1 was defined first. Determinations were based on the most frequent pathogenic (MFP) variant, rather than the accumulated variant frequency, due to more robust results with the former method. Strict thresholds were balanced against the high weight of the benign frequency rules, including possible mild phenotypes. The BS1 threshold was determined to be MFP × 1.5, but not <0.1%. Values falling between the MFP and the threshold were not assigned any rule. The thresholds for PM1 and BA1 were then set as one order of magnitude lower and higher, respectively, than the MFP, but not <0.5%. Frequencies between the PM1 and BS1 thresholds were assigned PM1_supporting. To avoid pathogenic variants with a very high frequency compared with others such as GJB2:p.Val37Ile having undue influence, and consistent with the recent ClinGen BA1 recommendations (https://www.onlinelibrary.wiley.com/doi/10.1002/humu.23642), the aiVCE algorithm automatically flagged outlier variants for placement on an exclusion list. Similarly, the BS2 threshold was determined using the minimum of the highest number of homozygous individuals with a pathogenic variant, plus a small delta.

Results: Benchmarking results indicate a high degree of concordance between the EP and aiVCE frequency thresholds, with an average difference <0.1% for the BS1 and PM1 rules that discriminate between pathogenic and benign classifications, and 0.08% between the MFP and the EP BS1 rule threshold. Most instances of discordance were due to the aiVCE employing a less strict threshold. A more significant difference in thresholds was observed in the PAH gene, which has mild effect phenotypes. The aiVCE thresholds for the most common variants, e.g., p.Ala300Ser (allele frequency 0.6%), p.Ala403Val (allele frequency 0.5%,), both of which result in partially reduced enzyme activity, and others, were significantly higher than the EP recommended threshold of 0.2% for BS1 (Figure). The final classification showed a concordance of 75.6% (227/300) between B/LB (benign/likely benign), VUS (variant of uncertain significance), P/LP (pathogenic/likely pathogenic), and 83% (249/300) when a VUS-subclassification was included (Table). These results were achieved without incorporating external data such as case-related evidences or functional studies extracted from the literature, commonly used by the EP groups.

Conclusions: Although an AI-based, gene-focused algorithm cannot replace EP recommendations, it can be used to facilitate and scale the process of implementing gene-specific recommendations.

Introduction

Gene-specific Frequency Criteria (PM2, BS1, BS2, BA1) in the American College of Medical Genetics and Genomics (ACMG)/Association for Molecular Pathology (AMP) Sequence Variant Interpretation Guidelines[1]

- The 2015 ACMG/AMP sequence variant interpretation guidelines provide a framework for classifying variants based on several benign (B) and pathogenic (P) evidence criteria
  - Includes criteria regarding the variant frequency and appearance in the general population versus patient population
- Gene-specific variant classification – employs thresholds specific to the gene; characteristics associated with the disease are crucial for accurate results
- Artificial intelligence (AI) could aid in the labor-intensive statistical calculations necessary for classification

Clinical Genome Resource (ClinGen) Expert Gene/Disease Panels (EPs)[2]

- EPs tasked with defining application of ACMG/AMP guidelines in specific genes/diseases
- Recently, the FDA recognized the genetic variant information in the Clinical Genome Resource (ClinGen) consortium's ClinGen Expert Curated Human Genetic Data as a source of valid scientific evidence that can be used to support clinical validity[3]

Artificial Intelligence-Based Variant Classification Engine (aiVCE)

- Data-driven; based on the ACMG/AMP sequence variant interpretation guidelines
- Automates majority of ACMG/AMP classification rules
- Variant classification accomplished by building prediction models at the gene and rule levels, based on various data sources (e.g., ClinVar, ClinGen, Uniprot, gnomAD, ExAC, Orphanet, etc.)
- Classification takes into account the gene and diseases associated with the variant
- Professional expertise can be applied to algorithm to determine thresholds specific to the gene being interrogated

Aims

- Assess the aiVCE's ability to predict refined thresholds for the frequency criteria to attain truly gene-specific classification
- Benchmark the aiVCE for implementation of FDA-recognized ClinGen EP gene-specific variant classification recommendations

Methods

aiVCE

- Developed and trained similar to the curation processes used by the different EPs for defining frequency-related rules (PM2, BS1, BS2, BA1)
- P variants extracted from ClinVar
- Corresponding population frequencies derived from the Genome Aggregation Database (gnomAD)

Published Gene-specific Datasets

- aiVCE assessed using recently published gene-specific datasets employed to create the recommendations put forth by each EP, both for frequency rule threshold determinations and variant classification benchmarking
  - CDH1[4]
  - PAH[7]
  - PTEN[5]
  - MYH[8]
  - Hearing loss[6]
  - RASopathy[9]

Frequency Rule Threshold Determinations

- Thresholds set to not miss P variants
- **BS1** – Allele frequency is greater than expected for disorder
  - Determined first (consistent with EPs)
  - Comparing multiple models for predicting this threshold showed that most frequent pathogenic (MFP) variants for each gene, along with P certainty level, provided most robust results in the training set experiments
  - Specific to this experiment, minimal hard thresholds (e.g., BS1 threshold ≥0.1%) were balanced against the high weight of B frequency rules, including not missing possible mild phenotypes
  - *BS1-Predicted* – The BS1 threshold was predicted according to a model based on MFP variants and the observed frequencies in each subpopulation, but not <0.1%.
  - *Values falling between the MFP and the threshold were not assigned any rule*
- **PM2 (Extremely low frequency in population databases) and BA1 (Allele frequency is >5% in ≥1 population database)**
  - Consistent with a number of EPs, set as one order of magnitude lower and higher, respectively, than the predicted threshold for BS1.
  - *Frequencies between PM2 and BS1 thresholds assigned PM2-supporting*
  - To avoid P variants with a very high frequency versus variants such as GJB2:p.Val37Ile having undue influence, and consistent with the recent ClinGen BA1 recommendations,[10] the aiVCE algorithm automatically flagged outlier variants for exclusion
- **BS2**
  - Determined based on the highest number per gene of homozygous individuals with a P variant

aiVCE vs. EP Variant Classification Benchmarking

- Primary analysis: P+LP versus B+LB+VUS
- Secondary analysis: With aiVCE subclassification of VUS to include VUS-Leaning P (VUS-LP) and VUS-Leaning B (VUS-LB)
- Due to lack of access, external data (e.g., case-related evidences or functional studies extracted from the literature) were not included by aiVCE, while external data are commonly employed by EPs

Results

Frequency Rule Threshold Determinations (Figures 1 and 2)

- Average differences between aiVCE and EPs <0.1% for the BS1 and PM2 rules that discriminate between P and B classifications (**Figure 1**)

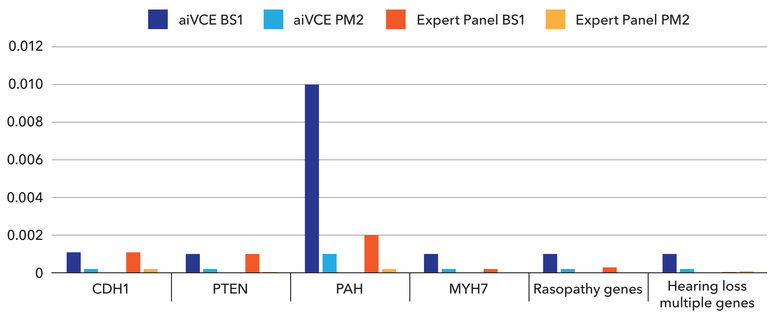

Figure 1. BS1 and PM2 rule thresholds: aiVCE versus EP

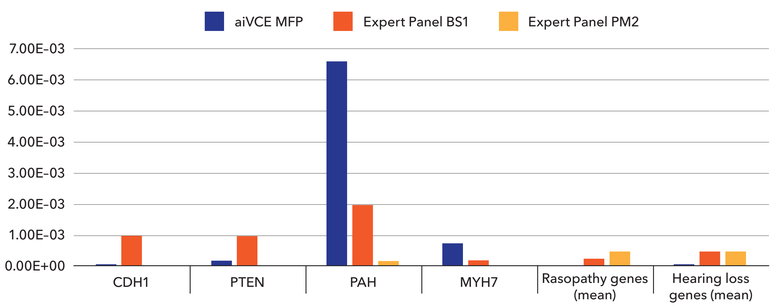

Figure 2. BS1 and PM2 rule thresholds: aiVCE BS1-Predicted variants versus EP

- Average difference of 0.08% between the aiVCE MFP variants and EP BS1 rule threshold (**Figure 2**)
- More of a difference in thresholds was observed in PAH gene, which has mild effect phenotypes
- The aiVCE thresholds for the most common variants in the PAH gene, e.g., p.Ala300Ser (allele frequency 0.6%), p.Ala403Val (allele frequency 0.5%), both of which result in partially reduced enzyme activity, and others, were significantly higher than the EP recommended threshold of 0.2% for BS1 (**Figure 2**)
- The variant scientist can adjust the threshold based on the desired level of phenotype severity

aiVCE vs. EP Variant Classification Benchmarking (Table 1)

- **Concordance**
  - Overall classification: 75.6% (227/300)
  - VUS-subclassification employed: 83.0% (249/300)
- **Discordance**
  - Most instances of discordance were due to external data available to EPs but not included in aiVCE algorithm

Table 1. aiVCE versus EP gene-specific variant classification recommendations

aiVCE EPs	B	LB	VUS	LP	P
B	29	1	1	0	0
LB	17	2	2	0	0
VUS-LB	12	3	6	1	1
VUS	3	0	29	8	11
VUS-LP	0	1	14	7	20
LP	0	0	3	29	95
P	0	0	0	0	5

CONCLUSIONS

- Setting proper thresholds, at the gene level, for ACMG criteria is important for accurate classification
- aiVCE demonstrated high concordance with the EP threshold, with an average difference of only 0.0008 (0.8%)
- Usage of very low thresholds can result in inappropriate application of the criteria for mild phenotypes in a gene
- Although an AI-based, gene-focused algorithm cannot supplant EPs, the aiVCE can be used to facilitate and scale the process of implementing gene-specific recommendation

References

1. Richards 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* 2015;17:405-24.
2. Rivera-Muñoz EA, e al. ClinGen Variant Curation Expert Panel experiences and standardized processes for disease and gene-level specification of the ACMG/AMP guidelines for sequence variant interpretation. *Hum Mutat* 2018;39:1614-22.
3. Lee K, et al. Specifications of the ACMG/AMP variant curation guidelines for the analysis of germline CDH1 sequence variants. *Hum Mutat* 2018;39:1553-68.
4. United States Food and Drug Administration. FDA takes new action to advance the development of reliable and beneficial genetic tests that can improve patient care. December 4, 2018 https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm627555.htm
5. Mester JL, et al. Gene-specific criteria for PTEN variant curation: Recommendations from the ClinGen PTEN Expert Panel. *Hum Mutat* 2018;39:1581-92.
6. Oza AM, et al. Expert specification of the ACMG/AMP variant interpretation guidelines for genetic hearing loss. *Hum Mutat* 2018;39:1593-613.
7. Zastrow DB, et al. Unique aspects of sequence variant interpretation for inborn errors of metabolism (IEM): The ClinGen IEM Working Group and the Phenylalanine Hydroxylase Gene. *Hum Mutat* 2018;39:1569-80.
8. Kelly MA, et al. Adaptation and validation of the ACMG/AMP variant classification framework for MYH7-associated inherited cardiomyopathies: recommendations by ClinGen's Inherited Cardiomyopathy Expert Panel. *Genet Med* 2018;20:351-9.
9. Gelb BD, et al. ClinGen's RASopathy Expert Panel consensus methods for variant interpretation. *Genet Med* 2018;20:1334-45.
10. Ghosh R, et al. Updated recommendation for the benign stand-alone ACMG/AMP criterion. *Hum Mutat* 2018;39:1525-30.