Gene-specific artificial intelligence-based variant classification engine: results of a time-capsule experiment

Yaron Einhorn (yaron@genoox.com)
Genoox
Moshe Einhorn
Genoox
Adaia Kamshov
Genoox
Oron Lev
Genoox
Amir Trabelsi
Genoox
Nurit Paz-Yaacov
Genoox
Susan J Gross
Icahn School of Medicine at Mount Sinai

Software

Keywords: variant, classification, ClinVar, artificial intelligence, algorithm

Posted Date: July 23rd, 2019

DOI: https://doi.org/10.21203/rs.2.11834/v1

Abstract

Background: Interpretation of genetic variation remains an impediment to cost-effective application of genomics to medicine. An advanced artificial intelligence (AI)-based Variant Classification Engine (aiVCE), rooted in ACMG/AMP guidelines, employs data-driven methods to expedite gene-specific classification (franklin.genoox.com). In this blinded study, the aiVCE's overall and rule-level performances were evaluated using ClinVar (v. 2018-10) variants with creation dates after 5/01/2017. By removing any prior knowledge of these variants from the aiVCE training data, they were treated as novel variants. Using a 'Full' dataset (75,801 variants with ≥1 star) and an 'Increased-Certainty' dataset (3,993 variants with ≥2 stars), the aiVCE classified variants as pathogenic (P), likely-pathogenic (LP), uncertain significance (VUS), likely-benign (LB), or benign (B). VUS with sufficient supporting data were subclassified as VUS-leaning benign or VUS-leaning pathogenic. aiVCE results were evaluated to determine concordance with final ClinVar classification and rule-level determinations. Results: The aiVCE demonstrated >97% concordance among Increased-Certainty variants. Concordance was >95% across variant effects (e.g., missense, null, splice region), and was >93.5% for the Full dataset. When assessing the aiVCE's application of specific ACMG rules, significant differences were observed between ClinVar P/LP and B/LB variants rule-met proportions (all P<0.00001), thus supporting gene-specific rule selections. Evaluation of discordance between the aiVCE and ClinVar uncovered evidences that might have been unavailable to submitting laboratories, highlighting AI utility in variant classification. Conclusions: The aiVCE exhibited robust performance, despite lacking past evidence, in determining whether variants would be categorized as P/LP. Applying latest computational advances to existing guidelines may assist scientists and clinicians interpret variants with limited clinical information and greatly reduce analytical bottlenecks.

Background

The science of human genomics has greatly benefited from the advent of high-throughput next-generation sequencing (NGS) technology. Information derived from the classification of variants is critical to discovering or confirming disease etiologies and guiding treatment guidelines and patient-specific plans. Variant classification, however, is a complex task, requiring assemblage and assessment of currently available information.[1]

As the use of NGS and shared archives has expanded, so has the volume of variant classification data and the need for novel analytical approaches.[2] In 2015, the American College of Medical Genetics and Genomics (ACMG) and Association for Molecular Pathology (AMP) jointly published standards and guidelines for variant classification to homogenize methods and reduce discordance between clinical laboratories.[3] These guidelines apply weighted rules across multiple categories such as variant frequency, variant type, association to previous reports for pathogenicity, and consistency with inheritance model. Subsequently, the Clinical Genome Resource (ClinGen) Expert Gene/Disease Panels (EPs) were tasked with defining application of ACMG/AMP guidelines in specific genes/diseases.[4]

The ACMG/AMP guidelines have been adopted worldwide. While the guidelines provide important direction to clinical and research geneticists, there remains a great deal of evidence data to assimilate into the classification process. Further, implementation of new recommendations for more accurate interpretation of variants involves demanding and complex bioinformatics work. As such, artificial intelligence (AI)-based solutions can be used to facilitate and scale the process of implementing gene-specific recommendations, enhancing current interpretation guidelines.

A novel AI-based Variant Classification Engine (aiVCE) algorithm has been developed to integrate knowledge from available databases and published literature. The aiVCE is data-driven, extracting existing evidential data from various sources on an ongoing basis, allowing for accurate, consistent, and rapid variant classification per ACMG/AMP standards and guidelines. The aiVCE algorithm can assess all classification criteria amenable to automation, while case level information, e.g., *de novo* evidence, segregation data, or allelic data, can be provided as additional input by the user. The aiVCE places particular emphasis on considering gene-specific evidence at the gene level, consistent with the latest efforts by ClinGen EPs.[5–11] Further, frequency-related rule thresholds for different variants of a specific gene can be customized.

Recent guidelines set forth by the AMP and College of American Pathologists emphasize the importance of validating pipeline tools and algorithms.[12] As such, we sought to validate the novel aiVCE using the ClinVar database[13] and conducted a blinded time-capsule experiment to predict the ability of our algorithm to classify variants that were only uploaded to ClinVar after the time-capsule cutoff date.

Implementation

Artificial intelligence-based variant classification engine (aiVCE)

The proprietary aiVCE is based upon ACMG/AMP standards and guidelines.[3] While these guidelines provide a clear and detailed framework for variant classification, implementation can vary. The aiVCE employs a data-driven, machine-learning process to implement variant classification rules for novel, as well as previously reported, variants. Specifically, the aiVCE continually assimilates information at the variant level from ClinVar, Uniprot, gnomAD, other public data sources, and in-house manually-curated variant databases to establish variant-level knowledge that is extrapolated to the gene level to propose a variant classification of benign (*B),* likely benign (*LB),* variant of uncertain significance (*VUS),*, likely pathogenic (*LP),* and pathogenic (*P).* For users desiring additional information about *VUS,* the aiVCE has the capability to further classify such variants into three categories: *VUS*-leaning benign *(VUS-LB), VUS,* and *VUS*-leaning pathogenic *(VUS-LP).* Specifically, the *VUS-LB* and *VUS-LP* variants would be those considered *VUS* according to the ACMG/AMP guidelines, but that have additional evidence available for consideration (i.e., the aiVCE would subclassify *VUS* as *VUS-LB* for variants with evidence for being *B* but without enough support for being classified as *LB,* or as *VUS-LP* for variants with evidence of being *P* yet

not enough for being classified as *LP).* The aiVCE automates 17/28 ACMG rules: PVS1, PS1, PM1, PM2, PM4, PM5, PP2, PP3, PP5, BA1, BS1, BS2, BP1, BP3, BP4, BP6, and BP7. The remaining rules cannot be automated, as they require clinical information specific to the patient genotype, e.g., familial data, *de novo* evidence (PS2, PM6), segregation data (PP1, BS4), and/or allelic data (PM3, BP2). The evidence for rules that cannot be automated, however, can be provided by the user as an input for the aiVCE. In addition, the aiVCE provides the option to override automated criteria. Gene symbols were derived from the HUGO Gene Nomenclature Committee database.[14] Transcript and exon information was based on the Reference Sequence (RefSeq) database.[15]

To moderate the propensity for discordance and conflicting interpretation inherent in variant classification,[1] we developed a model for assigning a confidence level to variant classification (Confidence Model). Specifically, an in-house curated dataset of thousands of variants, for which classification was manually curated and thus highly confident, was employed to identify variant features, e.g., number of submissions, submitting organization, conflicting evidence, and date of submission, contributing to variant classification certainty. The Confidence Model assigned a Confidence Score to submitted variants consolidated from ClinVar, UniProt, and other sources to create an internal Variant KnowledgeBase (VarKB), which forms the basis of classifying any new variant in the aiVCE. Further, VarKB data were employed to generate aggregated models at the gene level, i.e., a Gene KnowledgeBase (geneKB), to determine frequency thresholds and known disease mechanisms for the gene (*Figure 1).*.

Frequency threshold determinations

With a focus on not missing any *P* variants, the aiVCE emphasizes robust sensitivity in detecting *P* variants, even at the expense of possibly lowering specificity (i.e., slightly more falsely *P* determinations). For instance, in some cases where confidence of the evidence is suboptimal, the aiVCE would still consider a *P*-supporting rule as met, but could assign a lower strength or Confidence Score. Even in cases where statistical modeling or an AI approach determines whether a rule was met, the aiVCE provides comprehensive annotations for the clinician's use when determining the final classification. For example, if the PM1 rule (hotspot region) is met, the aiVCE would provide the numbers and examples of *P* variants found in that region.

Utilizing geneKB to determine frequency thresholds for rules related to Population Data (PM2, BA1, BS1, BS2), consistent with recent ClinGen EP methods,[4] that for BS1 (Allele frequency is greater than expected for disorder) was defined first. Comparing multiple models for predicting this threshold showed that most frequent pathogenic (MFP) variants for each gene, coupled with the *P* Confidence Score, provided the most robust results. Specific to this experiment, strict thresholds were balanced against the high weight of the *B*-supporting frequency rules, including possible mild phenotypes. The BS1 threshold was predicted using MFP variants and the observed frequencies in each subpopulation, but not <0.1%. The thresholds for PM2 and BA1 were then set as one order of magnitude lower and higher, respectively, than the predicted threshold for BS1, but not <0.5%. To avoid *P* variants with much higher frequency than

others (e.g., GJB2:p.Val37Ile) having undue influence, and consistent with the recent ClinGen EP BA1 recommendations,[5] the aiVCE algorithm automatically flagged outlier variants for placement on an exclusion list. Similarly, the BS2 threshold was predicted using the highest number of homozygous individuals with a *P* variant and their distribution among different subpopulations.

Classification rules related to variant effect on derivative protein include PVS1 (criterion for predicted loss-of-function [LOF] variants), PS1, PM5, PM4, BP1, BP3, and BP7. Recently, the ClinGen Sequence Variant Interpretation Workgroup published recommendations for more accurate interpretation of null variants.[16] These recommendations integrate a variety of evidences, including but not limited to, the functional role of the affected region, region size, variant effect, and variant location on the transcript, to develop a decision tree for assigning a final weight of evidence for calling the PVS1 rule, rather than relying on PVS strength only. The aiVCE model incorporates the majority of these recommendations into the classification algorithm. For the PS1/PM5 rules (same/different amino acid change previously established based on *P* variants in the 'Training' dataset), scores reflected the level of confidence in the source of the reported variant. For the BP3 and PM4 rules, the repetitive nature of the region was determined via RepeatMasker (http://www.repeatmasker.org/). Throughout, the effect of the variant on the protein was determined using RefSeq database transcripts.[15]

Specific to hotspot regions (PM1 rule) for each exon/domain, a sliding window within the aiVCE initially extracts candidate regions between each pair of *B* variants or at the edges of the region to be clear of *B* variants; candidate regions without *P* variants are ignored. Within each resulting candidate region, the aiVCE further detects inner borders of *P* variants contained within, and determines the number of *P* variants. Based on the density and overall number of *P* variants, the aiVCE evaluates each region for the presence of hotspots and then assigns a weight to the PM1 rule as 'supporting,' 'moderate,' or 'strong.' The aiVCE's weighting algorithm differentiates between the inner *P* region and the region between *P* and *B* variants.

For rules related to *in silico* predictions (PP3, BP4), we developed an in-house training dataset (based on VarKB), rather than relying on external algorithms trained on different data, to enhance prediction capabilities. Specifically, training data comprised *P* and *B* variants, excluding those previously used to train constituent tools. Employing a logistic regression algorithm to determine if "multiple lines of computational evidence support a deleterious effect on the gene or gene product," a single aggregated score was generated to determine if the variant was to be considered deleterious.

Application of the BP1 rule was determined based on the number of null *P* variants and the number of non-*P* variants in the gene, as well as their ratio. Similarly, calling of the PP2 rule was determined based on the number of missense *P* variants and missense *B* variants in the gene and their ratio.

Recent opinions suggest that the use of reputable source data rules (PP5, BP6) are preferable to expert opinion in the absence of primary data.[17] Given that primary data sharing remains a challenge, the PP5 and BP6 rules contribute to the aiVCE's ability to classify variants for which evidence, but not primary

data, exist. Note that in the experiments described in this paper, these rules were disabled, as the purpose of the time-capsule experiment was to assess the aiVCE's classification of a novel variant with no previously published data. Reputable sources were considered those used for creating varKB. Rule strength was adjusted according to the Confidence Score, as described previously.

Benchmarking experiments

We benchmarked the aiVCE classification model using ClinVar[13] (version 2018–10) variants with Reference/Submission accession version creation dates after May 1, 2017, representing 1.5 years of submissions with no overlap with the 'Train' dataset. Our experiments employed a 'Full' dataset, defined as variants with ≥1 star from Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory submitters without conflicts. As well, because the ClinVar database does not represent a gold-standard database, a subset of "Increased-Certainty" variants, i.e., variants with ≥2 stars from ≥2 submitters with no conflicts or that were EP-reviewed (https://www.ncbi.nlm.nih.gov/clinvar/docs/details/), was also interrogated. Further, evidence from clinical databases specific to the 'Test' set was excluded from the algorithm during training so as not to bias the results.

Concordance between the aiVCE and ClinVar variant classifications was assessed according to the five ACMG categories (*B, LB, VUS, LP, P)*,, as well using a medically relevant two-tier classification, i.e., *P/LP* vs. *B/LB/VUS.*[18–20] The percent concordances and accompanying 95% confidence intervals, determined using Wilson's method, are reported. As *VUS* classification does not represent the actual behavior of a variant, but only that its true classification is unknown, we also assessed concordance when excluding ClinVar's *VUS* (i.e., *P/LP* vs. *B/LB)*.. To evaluate the impact of the aiVCE's *VUS* subclassification system on final classification, we further examined how many of the ClinVar *VUS* variants were prioritized correctly. A 2-tailed-z-test was used to compute P-values when comparing observed proportions.

Concordance between the aiVCE and ClinVar classifications was also assessed by variant effect (e.g., missense, null, splice region), noting that each region is distinct, with no overlap. As such, canonical splice sites (± 1–2) are not included in splice regions (± 3–10), and intronic regions do not contain either of these regions. We further examined specific variants with conflicting classifications (i.e., discordance), to delineate mechanisms of differentiation between ClinVar and the aiVCE, and summarized the distribution of ClinVar variants according to the aiVCE application of ACMG rules.

As frequency-related rules pose a challenge owing to the different disease mechanisms, we assessed aiVCE performance for such rules across diseases using the following six gene panels from the Genomics England PanelApp (https://panelapp.genomicsengland.co.uk/): RASopathies, Hereditary Ataxia, Familial Breast Cancer, Hereditary Neuropathy, Hearing Loss, and Confirmed Fanconi Anaemia or Bloom Syndrome (FA/BS).

Results

Datasets

Our experiments employed a 'Full' dataset and an 'Increased-Certainty' dataset. The full dataset comprised 75,801 variants with ≥1 star, limited to ≥1 CLIA-certified laboratory submitters and without conflicting interpretations, spanning 3,115 different genes. Given that ClinVar does not represent a gold-standard database, a subset dataset of Increased-Certainty variants was also interrogated. Comprising 3,993 variants with ≥2 stars, across 638 different genes, variants in the Increased-Certainty dataset had ≥2 submitters with no conflicts or were EP-reviewed (https://www.ncbi.nlm.nih.gov/clinvar/docs/details/).

Variant classification concordance

Variant classification results are presented according to the five categories suggested by the ACMG guidelines, i.e., *B, LB, VUS, LP,* and *P,* as well as with application of *VUS* subclassification (to *VUS-LB* or *VUS-LP)* when sufficient supporting data were available. We assessed concordance using a two-tier classification based on medical importance, i.e., *P/LP* vs. *B/LB/VUS,* and also when excluding ClinVar's *VUS* variants, as this category does not reflect the true classification of a variant.

When employing 2-tier classification, the aiVCE and ClinVar classifications were highly concordant, i.e., 97.29% (95% CI: 96.79%–97.79%) and 93.78% (95% CI: 93.61%–93.95%) for the Increased-Certainty and Full datasets, respectively (*Tables 1 and 2).*. Excluding variants classified as *VUS* in ClinVar did not strongly impact results, with 97.36% (95% CI: 96.70%–98.02%) and 93.26% (95% CI: 93.01%–93.51%) concordance in the Increased-Certainty and Full datasets, respectively.

Details of VUS subclassification discordance

The aiVCE's ability to further classify *VUS* variants as *VUS-LB* or *VUS-LP* provides prioritization that could prove useful for the variant scientist. Specifically, the *VUS-LB* and *VUS-LP* variants would be those considered *VUS* according to the ACMG/AMP guidelines, but that have additional evidence available for consideration. Of the 58,067 Full dataset variants classified as *VUS* by the aiVCE, 7,282 (12.5%) were subclassified as either *VUS-LP* or *VUS-LB.* Discordance was observed for 91/7,282 (1.2%) variants called as *VUS-LP* by the aiVCE but that were *LB* in ClinVar and only 4/7,282 (0.05%) variants called as *VUS-LB* by the aiVCE but that were *LP* in ClinVar (*Table 1).*. When assessed using the Increased-Certainty dataset, in which case 314/2,961 (10.6%) *VUS* variants were subclassified by the aiVCE, only 1/314 (0.3%) subclassified variants demonstrated discordance with ClinVar (i.e., *VUS-LP* by aiVCE but *LB* in ClinVar) (*Table 2).*.

Variant classification by effect

When categorized by variant effects, missense variants comprised the largest group in both datasets (Full: 47.07%, Increased-Certainty: 47.02%), followed by synonymous variants (Full: 23.18%, Increased-Certainty: 30.14%); null variants including frameshift, stop gain, and splice donor/acceptor (Full: 13.70%, Increased-Certainty: 11.94%); splice region variants (Full: 8.25%, Increased-Certainty: 7.28%); intronic not located in a splice region and untranslated region (UTR) variants (Full: 6.34%, Increased-Certainty: 2.62%); and non-frameshift indel variants (Full: 1.69%, Increased-Certainty: 0.97%) (*Tables 3 and 4)*.. As expected, the majority of the *P/LP* variants were null variants, the majority of *VUS* were missense variants, and the majority of *B/LB* variants were synonymous variants.

When concordance between the aiVCE and ClinVar classifications was assessed by variant effect, strong agreement was attained across all groupings, even in variants typically considered difficult to classify, e.g., missense and splice region variants. Specifically, respective levels of concordance in the Full and Increased-Certainty datasets were 85.74% and 99.51% for null variants, 90.27% and 95.30% for missense variants, 99.91% and 99.92% for synonymous variants, 98.90% and 100% for intronic not located in a splice region//UTR variants, and 96.58% and 97.59% for variants located in a splice region.

Variant discordance

In the Full dataset, 90.46% (21,914/24,223) of the ClinVar *LB* variants were classified as *VUS* by the aiVCE, and 21,103 (96.30%) of these variants were very rare (BA1/BS1/BS2 rules not applied) synonymous variants or variants located in non-coding regions, but not within the splice region. Given that additional case-specific evidence, based on the ACMG/AMP guidelines, would be required to classify such variants as *LB* rather than *VUS,* we further examined specific variants with conflicting classifications to delineate mechanisms of differentiation between ClinVar and the aiVCE. Utilizing the Increased-Certainty dataset to minimize the likelihood of misclassified variants and to limit the overall number of variants being assessed, we evaluated instances of discordance between the aiVCE and ClinVar using three tiers of classification (*P/LP, VUS, B/LB)*.. Reassuringly, no 'Strong' conflicts, defined as cases where a *B/LB* ClinVar variant was classified as *P/LP* by the aiVCE, or where a *P/LP* ClinVar variant was classified as *B/LB* by the aiVCE, were observed.

'Moderate' conflicts, defined as cases in which a ClinVar *P/LP* variant was classified as *VUS* by the aiVCE, or vice versa, were observed for 107/3,993 (2.67%) of Increased-Certainty variants. Specifically, 49 variants were classified as *LP* by the aiVCE but as *VUS* in ClinVar, and 58 variants were classified as *VUS* by the aiVCE but as *P/LP* in ClinVar (see *Additional file 1)*..

For the moderately conflicted variants classified as *LP* by the aiVCE but that were *VUS* in CinVar, 46/49 were missense variants and were categorized as *LP* due to a number of supporting evidences, including PM1 (46/46 variants), PM2 (46/46), PM5 (20/46), PP2 (41/46), and/or PP3 (43/46) (*Figure 2)*.. When the aiVCE evidences were considered against those provided in ClinVar, few instances of disagreement occurred for the PM2 (extremely low frequency in population databases) and PP3 (*in silico* predictions) rules. However, while the aiVCE detected these variants in mutational hotspots (PM1) in all cases, ClinVar

annotations typically did not include a PM1 assessment. Additionally, a different amino acid change of a known *P* variant (PM5) was detected by the aiVCE, but commonly not reported by the ClinVar submitters (20/46 variants). As an example, for the variant NM_000257.3:c.4807G>A (p.Ala1603Thr) in the MYH7 gene, the PM1 rule was met, as the aiVCE identified 64 *P/LP* variants in that region located in the Myosin_tail_1 domain without a *B* variant. PM5 was met, as the variant c.4807G>C (p.Ala1603Pro) was previously classified as *P* in ClinVar; however, the evidence (PM5) was not annotated as such by ClinVar submitters.

For the alternate scenario of moderately conflicted variants, i.e., when a variant was classified as *VUS* by the aiVCE but was *P/LP* in ClinVar, 48/58 variants were missense variants (n = 41) or variants located in splice region (n = 7) (*Additional file 1).*. Manual examination of several variants, for which detailed classification information was available in ClinVar, indicated these variants were classified as *P/LP* based on additional evidence that was not available to the aiVCE, including patient-level data extracted manually from the literature (e.g., clinical information from the affected patient, *de novo* variant, segregation data) (*Additional file 1).*. For example, the variant NM_004863.3:c.547C>T in SPTLC2 gene (p.Arg183Trp) was called as *P,* as it has been reported to segregate with autosomal dominant hereditary sensory and autonomic neuropathy type 1C in two families (https://www.ncbi.nlm.nih.gov/clinvar/variation/487224/). Given that this information is not available to, and associated rules are not applied automatically by, the aiVCE, the classification remained *VUS*. Of note, the aiVCE subclassified 34/58 variants as *VUS-LP,* suggesting a greater likelihood the variant is *P/LP.*

In addition, seven of the 58 moderately conflicted variants classified as *VUS* by the aiVCE but as *P/LP* in ClinVar were null variants. Five of these null variants were splice donor/acceptor variants called as PVS1_Moderate based on the recent ClinGen EP recommendations,[16] as the reading frame was not disrupted and the altered region was not known to be critical to protein function. For the remaining two null variants, although they were called as PVS1_Very Strong, their frequency was above their specific gene threshold to meet the PM2 rule. Specifically, for NM_012144.3:c.389–1G>C (p.Gly134Arg), the aiVCE suggested a threshold of 0.00111 for applying PM2 in the DNAI1 gene, yet due to a frequency of 0.00128 in the gnomAD (https://gnomad.broadinstitute.org) 'Other' population, the PM2 rule was not met. For NM_199292.2:c.457C>T (p.Arg153Ter), the suggested frequency for applying PM2 for the TH gene exceeded 0.0005, yet it appeared at a frequency of 0.0007 in the gnomAD East Asian population. The observed frequency for both variants was very close to their gene-specific suggested PM2 threshold. Although these variants did not meet the PM2 rule, they were still below suggested thresholds for applying BS1, owing to the fact that the aiVCE has an uncertain frequency region for determining PM2 or BS1 application in cases where no rule is applied.

Of particular note, five discordant variants, which were classified as *VUS* by the aiVCE and *LP* in ClinVar, occurred in MSH2 and MLH1 genes that were later (based on ClinVar[13] version 2019/1) reclassified in ClinVar by The International Society for Gastrointestinal Hereditary Tumors (https://www.insight-group.org/) from *LP* to *VUS* (https://www.ncbi.nlm.nih.gov/clinvar/23203244/).

Variants and ACMG rules

The distribution of ClinVar variants, according to the aiVCE application of ACMG rules and ClinVar submitter classifications, is shown in *Figure 3.* A significant difference (p<0.0001) was observed for use of *P*-supporting rules, as well as for application of *B*-supporting rules, to *P/LP* vs. *B/LB* variants.

When considering gene-specific rules for missense variants (*Figure 4a),* which are among the most difficult variants to classify, the aiVCE differentially applied *P*-supporting (PS1, PM1, PM5, PP2, PP3) rules to *P/LP* variants, and *B*-supporting (BP1) rules to *B/LB* variants (P<0.00001 for each rule). Further, application of the PVS1 rule was significantly different between *P/LP* vs. *B/LB* LOF variants (P<0.00001) (*Figure 4b)*..

The aiVCE effectively differentiated between *P/LP* vs. *B/LB* variants for both missense variants (*Figure 4a)* and variants located in splice regions (*Figure 4c)*.. The PP3 rule was applied for 62.18% of the missense and 76.19% of the splice region *P/LP* variants but for only 7.0% and 1.6% of the missense and splice region *B/LB* variants, respectively (P<0.00001). Similarly, the BP4 rule was applied for 60.7% and 66.7% of the *B/LB* missense and splice region variants, respectively, but for only 9.1% and 4.7% of the corresponding *P/LP* variants (P<0.00001). As expected, the BP7 rule was applied to all synonymous variants (*Figure 4d)*.. While B-supporting frequency rules were applied to more *B/LB* than *P/LP* variants, the difference was not significant.

When assessing the aiVCE's application of ACMG frequency-related rules (BA1, BS1, BS2, PM2), significant differences were observed between ClinVar *P/LP* vs. *B/LB* missense, LOF and splice region variants (P<0.00001 for each rule), thus supporting the gene-specific thresholds selected for these rules. Given that more rare (and thus difficult to classify) than common variants are submitted to ClinVar, it is not surprising that the aiVCE applied the PM2 rule to >50% of variants within each type of variant effect and for 99.67% and 83.18% of the ClinVar *P/LP* vs. *B/LB* variants, respectively (*Figure 4)*..

aiVCE classification performance across disease categories

To examine the robustness of the aiVCE across different disease groups, the following six gene panels from the Genomics England PanelApp (https://panelapp.genomicsengland.co.uk/) were employed: RASopathies, Hereditary Ataxia, Familial Breast Cancer, Hereditary Neuropathy, Hearing Loss, and Confirmed FA/BS. Utilizing only variants of genes in each panel from the Full dataset, concordance was consistently high across the disease categories evaluated, i.e., 92.67%–98.40% (*Table 5)*..

Also utilizing ClinVar *P/LP* variants of genes from each panel in the Full dataset, we assessed the performance of frequency-related rules across varying disease mechanisms. The PM2 rule (extremely low frequency in population databases) was met in >99.7% of the *P/LP* variants, the BS1 rule was not met or

met for only 1–2 (<0.1%) of the *P/LP* variants, and the BA1 rule was never met for *P/LP* variants (*Figure 5)*,, indicating that gene-specific and varied gene frequencies can be incorporated into the AI.

Specific to PM2 thresholds used for the genes in the different disease panels (*Figure 6)*,, RASopathy genes with a dominant inheritance model had a low threshold (mean [SD]: 0.0005 [0.0]), while genes with a recessive inheritance model and/or lower penetrance had much higher thresholds, e.g., FA/BS mean (SD): 0.0021 (0.0023) and Familial Breast Cancer mean (SD): 0.0019 (0.003). On average, the PM2 threshold was <0.0015 across disease categories.

When assessing the distribution of CinVar *P/LP* variants by aiVCE rule application, differences across disease categories were observed for application of missense/LOF-related rules (i.e., PVS1, PS1, PM1, PM5, PP2, PP3), as expected, owing to the different disease mechanisms represented by each gene panel. For example, while the PSV1 rule was met for 96.50% of the variants in the FA/BS panel and 93.11% of Familial Breast Cancer variants, it was met for only 72.00% of RASopathy, 80.45% of Hearing Loss, 76.87% of Hereditary Neuropathy, and 81.86% of Hereditary Ataxia variants. Conversely, the PS1, PM5, and PP2 rules were met for significantly higher proportions of RASopathy variants when compared with other gene groups (all P<0.00001), and the PM1 rule was met for 11.39% of Hereditary Neuropathy but only 47.69% of FA/BS variants (P<0.00001) (*Figure 5)*..

Discussion

Herein, we benchmarked an aiVCE algorithm, previously shown to be a robust platform for comprehensive downstream analysis and identification of DNA variants responsible for disease.[21–24] Specifically, we assessed aiVCE concordance with final ClinVar classifications and aiVCE rule-level determinations. Results reported herein suggest that the aiVCE has the potential to streamline variant classification for the variant scientist by automating ACMG rules for which supporting evidence is available. Despite the exclusion of ClinVar-derived data from the algorithm to guard against potential bias, and automating only some of the rules, the aiVCE demonstrated robust (>97%) concordance in determining whether future variants would be categorized as *P/LP.* Further, the aiVCE accurately predicted thresholds for variant/allele frequency-based rules. The aiVCE-determined PM2 (extremely low frequency) thresholds, which averaged <0.0015 across disease categories, were several orders of magnitude lower than the typical default hard threshold (0.005 or 0.010) suggested by other tools,[25,26] while BA1 (common allele) was not met (0%) and BS1 (frequency greater than expected) was met for <0.1% of the *P/LP* variants.

Observations related to the PP3 and BP4 rules suggest that the aiVCE's AI-based prediction model was sensitive and specific in classifying variants. For example, for splice region variants, the aiVCE called BP4 for 67.0% of the *B* variants and only 4.7% of the *P* variants, while the PP3 rule was met for 76.1% of the *P* variants and only 1.5% of the *B* variants. Application of other gene-specific rules, including PVS1, PM1, PP2, and BP1, also exhibited significant differentiation between *B/LB* and *P/LP* variants. These results

show the utility of the aiVCE in applying gene-specific evidence and knowledge that may prove useful to the variant scientist.

Many of the rare *B/LB* variants in ClinVar were classified as *VUS* by the aiVCE. Most of these variants are very rare non-coding region variants or synonymous variants not located near a splice region, for which an additional case-specific evidence is required for classification as *B/LB.* With the advent of deep splice/synonymous prediction tools,[27] the PP3/BP4 rules should be useful in providing evidence required to re-classify these *VUS* to *LB.* The ACMG/AMP criteria comprise more *P*-supporting than *B*-supporting rules, and inclusion of additional *B*-supporting evidence may be warranted. For example, similar to the PM5/PS1 rules, the same codon with the same/different amino acid change as a known *B* variant could be employed as *B*-supporting evidence, in line with the Evidence-based Network for Interpretation of Germline Mutant Alleles' (https://enigmaconsortium.org) classification criteria for BRCA1/2 genes. In the same way, as the PVS1 rule is applied to genes in which LOF is a common mechanism for disease, no *B*-supporting rule exists for null variants of genes in which LOF is known not to be a mechanism for a disease. Thus, these *B*-supporting evidences should be considered as well.

Examination of discordance between the aiVCE and ClinVar highlights the importance of performing re-analysis on a regular basis. Specifically, several variants classified as *VUS* by ClinVar submitters were found to be *LP* by the aiVCE, which might suggest evidence was unknown to the submitter at the time, e.g., different amino acid change than in a known *P* variant (PM5), or lack of clinical domain expert knowledge such as a gene's hotspot region or critical functional domain (PM1), further demonstrating the utility of the aiVCE. On the other hand, development of a more structured approach to incorporate case level information by databases such as ClinVar would provide additional evidence currently missing in the aiVCE, such as segregation or *de novo* data. In the current era of AI, such evidences can be rapidly assimilated into existing and in-development algorithms.

The aiVCE is a data-driven, AI-based tool that relies on previous evidence derived from various data sources, making data-sharing and community initiatives like ClinVar and gnomAD essential. The availability of more evidence for the aiVCE to train its models only improves its ability to provide accurate variant assessments. The time-capsule experiment illustrates how data sharing is critical to reducing uncertainty in variant classification, not only for the specific variant that was shared, but also for novel variants of the gene. Without any prior evidence, the aiVCE begins with a default threshold for the PM2 rule of 1%, and as more evidence is gathered, the threshold is further refined. Even evidence from only a few variants can impact thresholds, as evidenced by the frequency-related rule thresholds for the novel *P* variants NM_012144.3:c.389–1G>C and NM_199292.2:c.457C>T (p.Arg153*), which at the time of the time-capsule experiment were not known to be *P,* and as such were not considered as part of the model. Similarly, the increase in concordance from 85.74% of variants with ≥1 star (Full dataset) to 99.51% of variants with ≥2 stars (Increased-Certainty dataset) can be attributed to the greater certainty variants occurring in genes for which more evidence existed, yielding more accurate classification. The increased concordance can also be attributed to, among other features of the aiVCE tool, improved application of the PVS1 rule, which requires the aiVCE to correctly determine whether LOF is a known mechanism of

disease if *P* null variants exist in the gene, as well as whether a skipped/truncated region is critical to gene function.

As the aiVCE is data- and AI-driven at the gene level, the increased availability of evidence for a specific gene improves the aiVCE's accuracy. Conversely, the aiVCE's utility for unknown genes is constrained for application of several rules, as well as for determining more global thresholds. Such limitations also exist for clinical laboratories evaluating genes with less available clinical information and, thus, are not unique to the aiVCE.

Conclusions

Our aiVCE, even without input from clinical databases specific to the 'Test' set, could predict with very high concordance whether a variant in the future would be categorized as *P/LP.* Robust platforms that facilitate comprehensive downstream analysis by applying the latest computational methodologies to ACMG/AMP guidelines may assist variant scientists with classification and interpretation of variants, including those with limited clinical information.

Availability And Requirements

Project name: aiVCE

Project home page: franklin.genoox.com

Archived version: Not applicable

Operating system(s): Platform independent

Programming language: Web

Other requirements: None

License: The tool is offered as a completely free service to the community under franklin.genoox.com.

Any restrictions to use by non-academics: None

Abbreviations

ACMG - American College of Medical Genetics and Genomics

aiVCE—artificial intelligence-based variant classification engine

AMP - Association for Molecular Pathology

*B*benign

BS—Bloom syndrome

CLIA - Clinical Laboratory Improvement Amendments

ClinGen - Clinical Genome Resource

EP - Expert Gene/Disease Panel

FA - Fanconi anaemia

geneKB - Gene KnowledgeBase

*LB*likely benign

MFP - most frequent pathogenic

LOF—loss of function

*LP*likely pathogenic

NGS - next-generation sequencing

P - pathogenic

RefSeq - Reference Sequence

VarKB - Variant KnowledgeBase

UTR - untranslated region

*VUS*variant of uncertain significance

Declarations

Ethics approval and consent to participate—Not applicable

*Consent for publication*Not applicable

Availability of data and material –The datasets generated and/or analyzed during the current study are available at: https://github.com/genoox/aivce/.

*Competing interests*Y Einhorn, M Einhorn, Adaia Kamshov, MD, Oron Lev, MSc, A Trabelsi, and N Paz-Yaacov are employed by Genoox. S Gross is a paid consultant of Genoox.

*Funding*Authors who are employed by or who consult for Genoox were involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Genoox provided

funding for a professional medical writer, Michelle L Perate, MS, to assist the authors with manuscript preparation and submission.

*Authors' contributions*Study design and conduct (YE, ME, AK, OL, AT, NP-Y, SG); Data collection, management, analysis (YE, AK), and interpretation (YE, ME, AK, OL, AT, NP-Y, SG); Manuscript preparation (YE, ME, SG), review (AK, OL, AT, NP-Y), approval (YE, ME, AK, OL, AT, NP-Y, SG); Decision to submit the manuscript for publication (YE, ME, AK, OL, AT, NP-Y, SG); Full access to all data and takes responsibility for integrity of the data and the accuracy of the data analysis (YE, ME, AK, OL, AT, NP-Y, SG)

Acknowledgements—The authors thank Elaine Lyon, PhD for manuscript review and Michelle L Perate, MS for assistance with manuscript preparation and submission.

References

1. Harrison SM, Dolinsky JS, Knight Johnson AE, Pesaran T, Azzariti DR, Bale S, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. Genet Med. 2017;19:1096–104.
2. Green ED, Rubin EM, Olson MV. The future of DNA sequencing. Nature. 2017;550:179–181.
3. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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.
4. Rivera-Muñoz EA, Milko LV, Harrison SM, Azzariti DR, Kurtz CL, Lee K, et 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.
5. Ghosh R, Harrison SM, Rehm HL, Plon SE, Biesecker LG; ClinGen Sequence Variant Interpretation Working Group. Updated recommendation for the benign stand-alone ACMG/AMP criterion. Hum Mutat. 2018; 39:1525–30.
6. Lee K, Krempely K, Roberts ME, Anderson MJ, Carneiro F, Chao E, et al. Specifications of the ACMG/AMP variant curation guidelines for the analysis of germline CDH1 sequence variants. Hum Mutat. 2018;39:1553–68.
7. Mester, JL, Ghosh R, Pesaran T, Huether R, Karam R, Hruska KS, et al. Gene-specific criteria for PTEN variant curation: Recommendations from the ClinGen PTEN Expert Panel. Hum Mutat. 2018;39:1581–92.
8. Oza AM, DiStefano MT, Hemphill SE, Cushman BJ, Grant AR, Siegert RK, et al. Expert specification of the ACMG/AMP variant interpretation guidelines for genetic hearing loss. Hum Mutat. 2018;39:1593–613.
9. Zastrow DB, Baudet H, Shen W, Thomas A, Si Y, Weaver MA, 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.

10. Kelly MA, Caleshu C, Morales A, Buchan J, Wolf Z, Harrison SM, 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.
11. Gelb BD, Cavé H, Dillon MW, Gripp KW, Lee JA, Mason-Suares H, et al. ClinGen's RASopathy Expert Panel consensus methods for variant interpretation. Genet Med. 2018; 20:1334–45.
12. Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. J Mol Diagn. 2018;20:4–27.
13. ClinVar. https://www.ncbi.nlm.nih.gov/clinvar/.
14. Yates B, Braschi B, Gray KA, Seal RL, Tweedie S, Bruford EA. Genenames.org: the HGNC and VGNC resources in 2017. Nucleic Acids Res. 2017;45:D619–25.
15. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44:D733–45.
16. Tayoun AN, Pesaran T, DiStefano MT, Oza A, Rehm HL, Biesecker LG, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat. 2018;39:1517–24.
17. Biesecker LG, Harrison SM, ClinGen Sequence Variant Interpretation Working Group. The ACMG/AMP reputable source criteria for the interpretation of sequence variants. Genet Med. 2018;20:1687–8.
18. Yang S, Lincoln SE, Kobayashi Y, Nykamp K, Nussbaum RL, Topper S. Sources of discordance among germ-line variant classifications in ClinVar. Genet Med. 2017;19:1118–26.
19. 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;35:1259–61.
20. Balmaña J, Digiovanni L, Gaddam P, Walsh MF, Joseph V, Stadler ZK, 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;34:4071–8.
21. Guemez-Gamboa A, Çağlayan AO, Stanley V, Gregor A, Zaki MS, Saleem SN, et al. Loss of protocadherin–12 leads to diencephalic-mesencephalic junction dysplasia syndrome. Ann Neurol. 2018;84:638–47.
22. Kurolap A, Eshach-Adiv O, Gonzaga-Jauregui C, Dolnikov K, Mory A, Paperna T, et al. Establishing the role of PLVAP in protein-losing enteropathy: a homozygous missense variant leads to an attenuated phenotype. J Med Genet. 2018;55:779–84.
23. Birnbaum R, Yosha-Orpaz N, Yanoov-Sharav M, Kidron D, Gur H, Yosovich K, et al. Prenatal and postnatal presentation of PRMT7 related syndrome: Expanding the phenotypic manifestations. Am J Med Genet A. 2019;179:78–84.

24. Kurolap A, Eshach-Adiv O, Hershkovitz T, Paperna T, Mory A, Oz-Levi D, et al. Loss of CD55 in eculizumab-responsive protein-losing enteropathy. N Engl J Med. 2017;377:87–9.
25. Ravichandran V, Shameer Z, Kemel Y, Walsh M, Cadoo K, Lipkin S, et al. Toward automation of germline variant curation in clinical cancer genetics. Genet Med. 2019 Feb 21. doi: 10.1038/s41436–019–0463–8. [Epub ahead of print] Erratum in: Genet Med. 2019 Mar 21; doi: 10.1038/s41436–019–0490–5.
26. LI Q, Wang K. InterVar: Clinical interpretation of genetic variants by the 2015 ACMG-AMP Guidelines. Am J Hum Genet. 2017;100:267–80.
27. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting splicing from primary sequence with deep learning. Cell. 2019;176:535–48.e24.

Tables

Table 1. Benchmarking an aiVCE using a time-capsule of the ClinVar database – Full dataset

	aiVCE	*B*	*LB*	*VUS-LB*	*VUS*	*VUS-LP*	*LP*	*P*
linVar								
3		2096	1093	203	806	9	**4**	**0**
.B		588	1719	865	20967	82	**32**	**0**
/US		61	1562	372	27807	4342	**2040**	**5**
.P		**5**	**5**	**3**	**919**	**1125**	4497	3
P		**5**	**3**	**1**	**286**	**280**	4007	9

aiVCE	*P/LP*	*B/LB/VUS*	**Concordance**
ClinVar			**(95% CI)**
P/LP	8516	**2632**	93.78%
B/LB/VUS	**2081**	62572	(96.70%-98.02%)

Note: Table cells exhibiting bolded numbers represent instances of potential discordance.

aiVCE, artificial intelligence-based Variant Classification Engine; *B*, benign; CI, confidence interval; *LB*, likely benign; *LP*, likely pathogenic; *P*, pathogenic; *VUS*, variant of uncertain significance; *VUS-LB*, *VUS-*

leaning benign; *VUS-LP*, *VUS*-leaning pathogenic

Table 2. Benchmarking an aiVCE using a time-capsule of the ClinVar database – Increased-Certainty dataset

aiVCE	*B*	*LB*	*VUS-LB*	*VUS*	*VUS-LP*	*LP*	*P*
ClinVar							
B	123	44	9	10	0	**0**	**0**
LB	24	146	63	1286	1	**1**	**0**
VUS	0	172	41	1327	166	**49**	**0**
LP	**0**	**0**	**0**	**12**	**16**	58	0
P	**0**	**0**	**0**	**12**	**18**	415	0

aiVCE	*P/LP*	*B/LB/VUS*	**Concordance (95% CI)**
ClinVar			
P/LP	473	**50**	97.29%
B/LB	**58**	3412	(96.79%-97.79%)

Note: Table cells exhibiting bolded numbers represent instances of potential discordance.

aiVCE, artificial intelligence-based Variant Classification Engine; *B*, benign; CI, confidence interval; *LB*, likely benign; *LP*, likely pathogenic; *P*, pathogenic; *VUS*, variant of uncertain significance; *VUS-LB*, *VUS*-leaning benign; *VUS-LP*, *VUS*-leaning pathogenic

Table 3. Distribution of variant effects overall and by ClinVar classification – Increased-Certainty dataset

Effects	All ClinVar	ClinVar *P*	ClinVar *VUS*	ClinVar *B*
Number of variants	3993	531	1755	1707
Null variants, n (%)	477 (11.94%)	453	19	4
Frameshift	295	287	8	0
Nonsense	119	110	7	2
Splice Donor/Acceptor	59	56	1	2
Start-loss	4	1	3	0
Intronic and UTR, n (%)	105 (2.62%)	0	5	100
Splice Region[a], n (%)	291 (7.28%)	7	63	221
Synonymous, n (%)	1204 (30.14%)	1	11	1192
Missense, n (%)	1878 (47.02%)	66	1627	184
Non-frameshift Indels, n (%)	39 (0.97%)	3	30	6

[a]Splice region defined as the region within ± 3-10 bases of the exon intron boundary, i.e., does not include splice acceptor/donor regions.

B, benign; *P*, pathogenic; UTR, untranslated region; *VUS*, variant of uncertain significance

Table 4. Distribution of variant effects overall and by ClinVar classification – Full dataset

Effects	All ClinVar	ClinVar *P*	ClinVar *VUS*	ClinVar *B*
Number of variants	75801	11148	36189	28464
Null variants, n (%)	1262 (13.7%)	8552	1465	87
Frameshift	3940	3369	540	31
Nonsense	3306	2793	494	19
Splice Donor/Acceptor	2601	2239	331	31
Start-loss	257	151	100	6
Intronic and UTR, n (%)	4812 (6.34%)	86	637	4089
Splice Region[a], n (%)	6256 (8.25%)	168	1944	4144
Synonymous, n (%)	17577 (23.18%)	11	1066	16500
Missense, n (%)	35684 (47.07%)	2224	29919	3541
Non-frameshift Indels, n (%)	1288 (1.69%)	94	1099	95

[a]Splice region defined as the region within ± 3-10 bases of the exon intron boundary, i.e., does not include splice acceptor/donor regions.

B, benign; *P*, pathogenic; UTR, untranslated region; *VUS*, variant of uncertain significance

Table 5. Concordance between the aiVCE and ClinVar classifications of *P/LP* variants by disease – Full dataset

Disease	Number of variants	Concordance[a]
Hereditary Neuropathy	11943	94.78%
RASopathies	1194	94.52%
Hereditary Ataxia	4553	94.24%
Familial Breast Cancer	7164	98.25%
Hearing Loss	7617	92.67%
Fanconi anemia/Bloom syndrome	3633	98.40%

[a]Based on 2-tier classification (*P/LP* vs. *B/LB/VUS*)

aiVCE, artificial intelligence-based Variant Classification Engine; *LP*, likely pathogenic; *P*, pathogenic

Figures

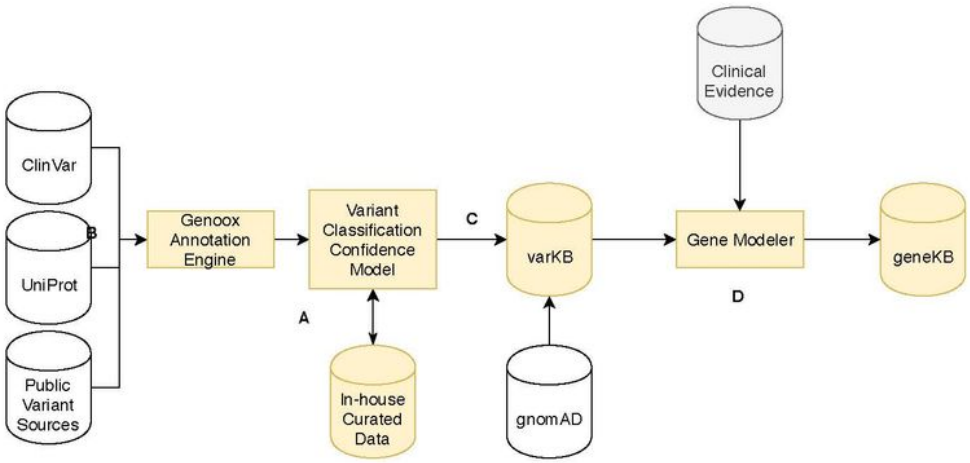

Figure 1

Building the training data for the aiVCE. In-house curated data were employed to determine features predicting variant classification confidence (Confidence Model; A); Using consolidated ClinVar, UniProt, and public source variants (B), the Confidence Model assigned a Confidence Score to the submitted classification to create an internal Variant KnowledgeBase (VarKB; C); Data in varKB were employed to create an internal Gene Knowledgebase (geneKB; D) of aggregated models at the gene level. artificial intelligence-based Variant Classification Engine

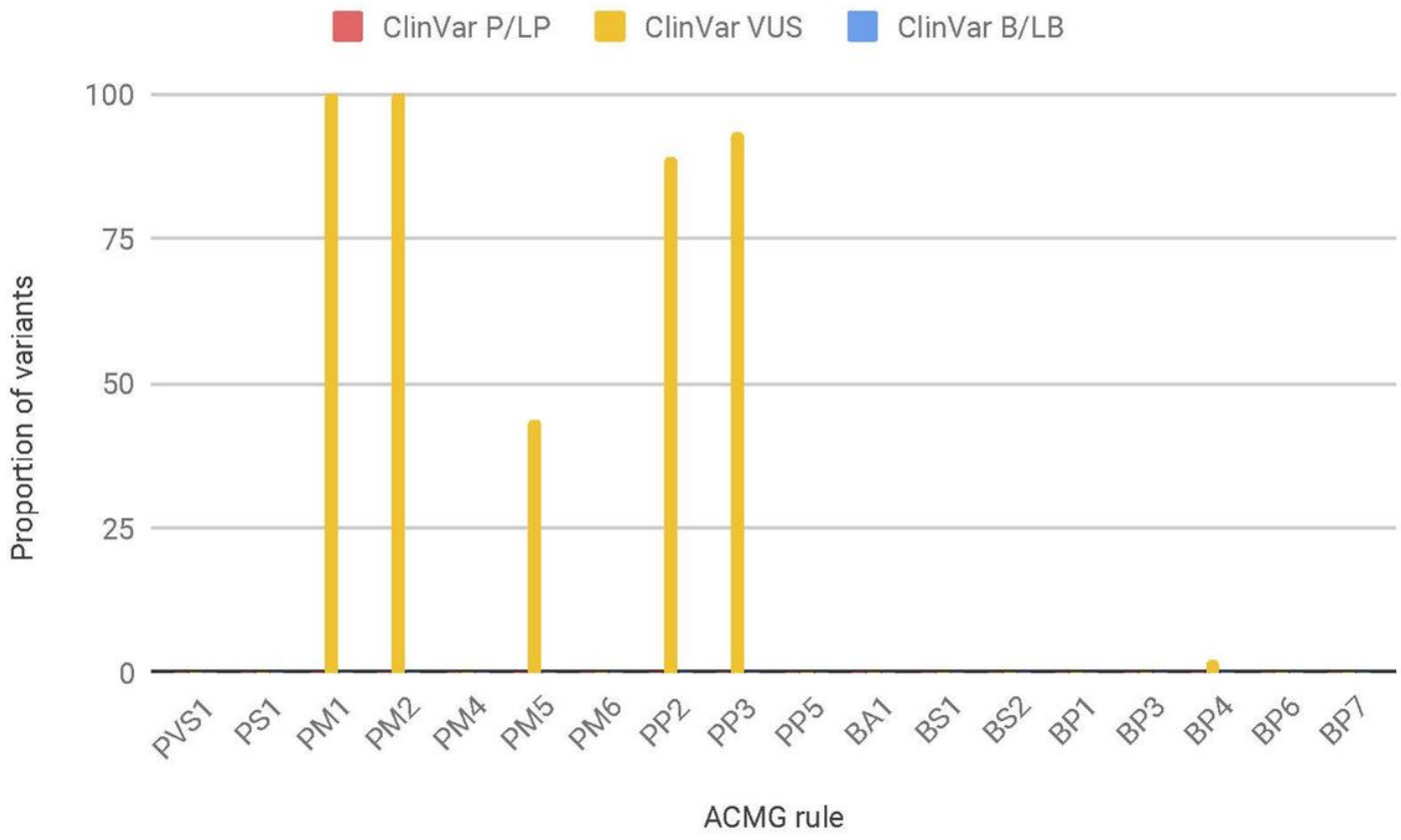

Figure 2

Distribution of ClinVar variants by aiVCE application of ACMG rules – Full dataset. Distribution is presented for 49 moderately discordant missense variants classified as LP by the aiVCE but as VUS in ClinVar; percentages derived as number of variants for which each rule is met divided by total number of variants each ClinVar classification (P/LP, VUS, B/LB). ACMG, American College of Medical Genetics and Genomics; aiVCE, artificial intelligence-based Variant Classification Engine; B, benign; LB, likely benign; LP, likely pathogenic; P, pathogenic; VUS, variant of uncertain significance

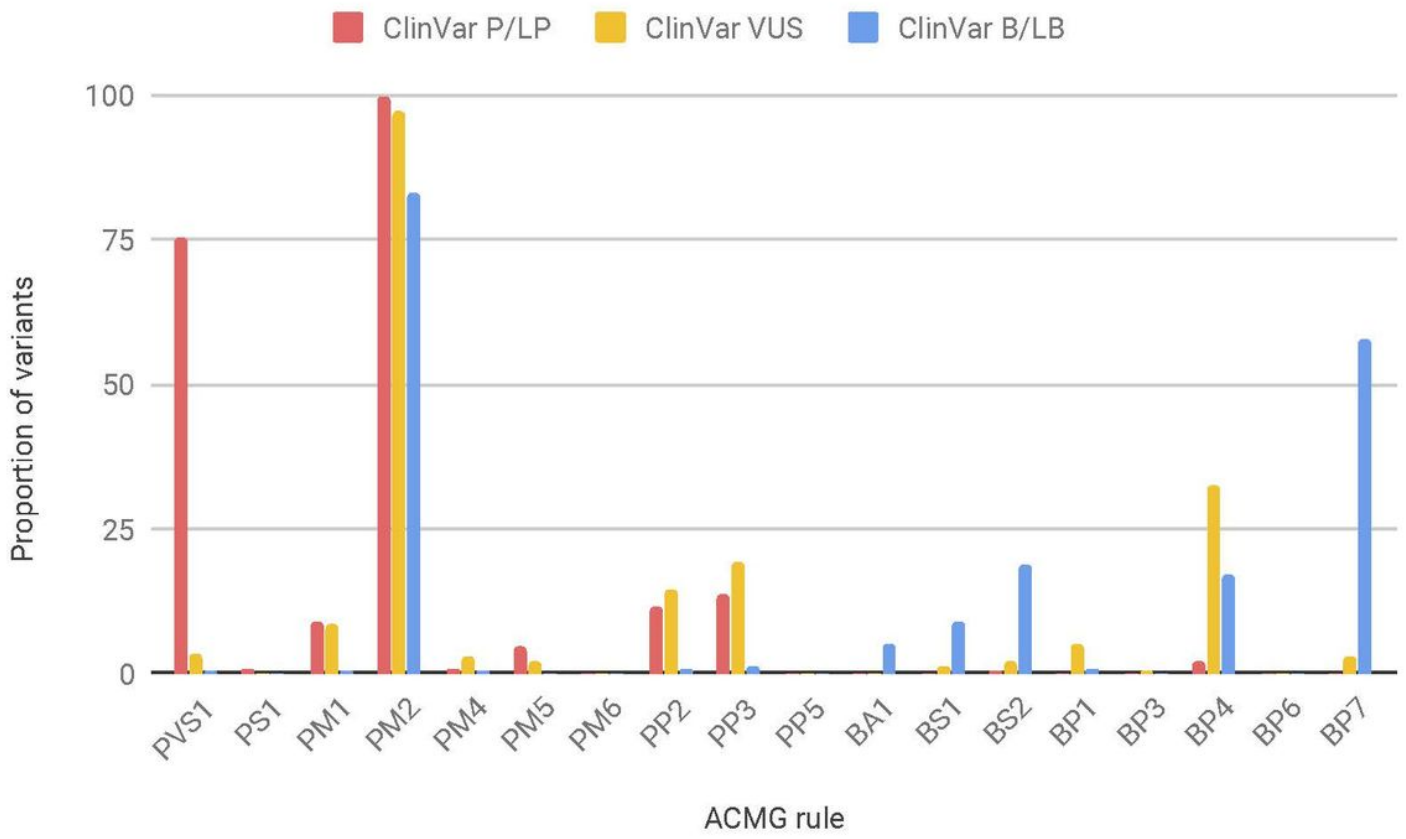

Figure 3

Distribution of ClinVar variants by aiVCE application of ACMG rules – Full dataset. Percentages derived as number of variants for which each rule is met divided by total number of variants each ClinVar classification (P/LP, VUS, B/LB). ACMG, American College of Medical Genetics and Genomics; aiVCE, artificial intelligence-based Variant Classification Engine; B, benign; LB, likely benign; LP, likely pathogenic; P, pathogenic; VUS, variant of uncertain significance

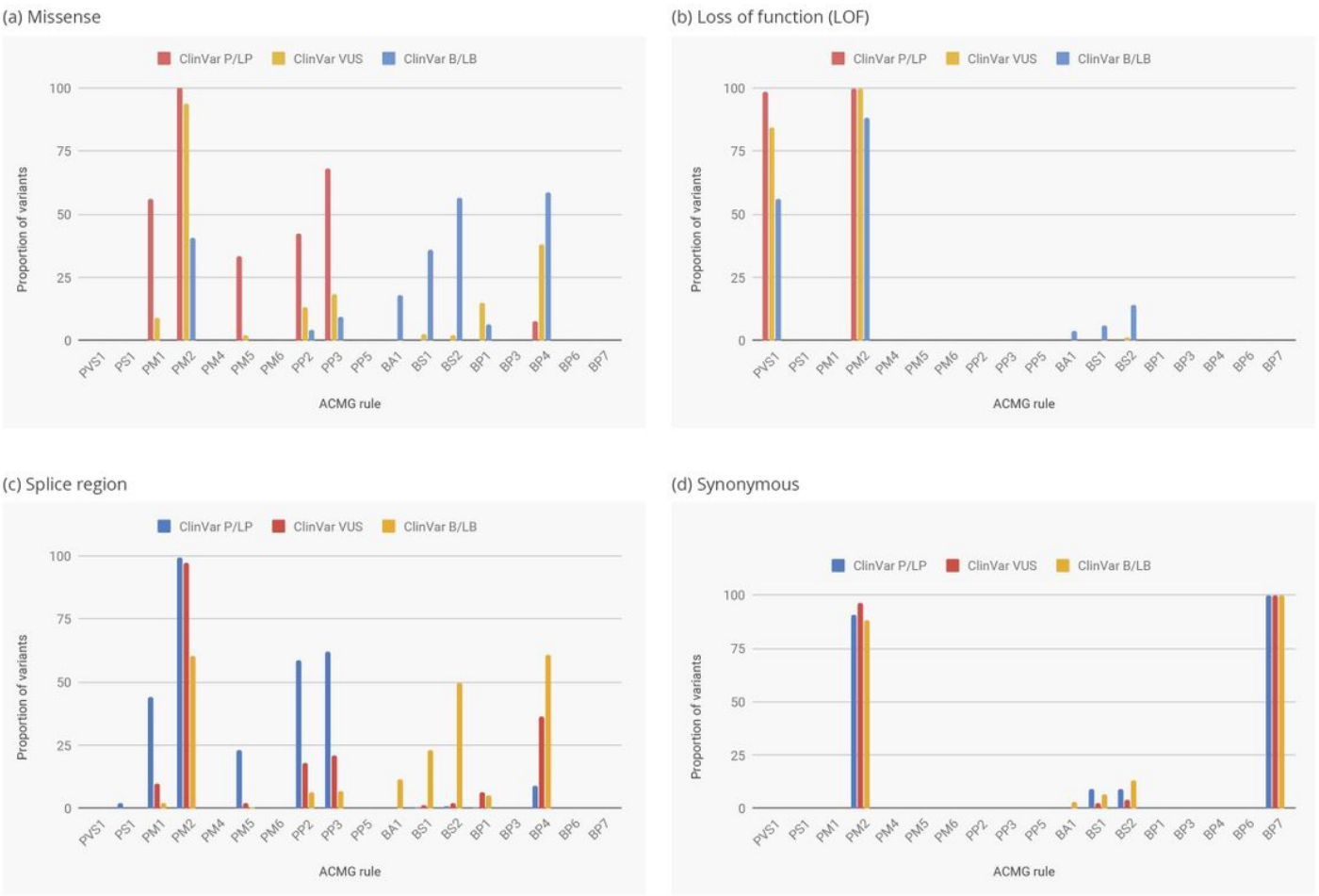

Figure 4

Distribution of ClinVar variants by aiVCE application of ACMG rules – Full dataset. Results shown for missense (a), loss of function (LOF; b), splice region (c), and synonymous (d) variant effects. Percentages were derived as number of variants for which each rule is met divided by total number of variants each ClinVar classification. ACMG, American College of Medical Genetics and Genomics; aiVCE, artificial intelligence-based Variant Classification Engine; B, benign; LB, likely benign; LP, likely pathogenic; P, pathogenic; VUS, variant of uncertain significance

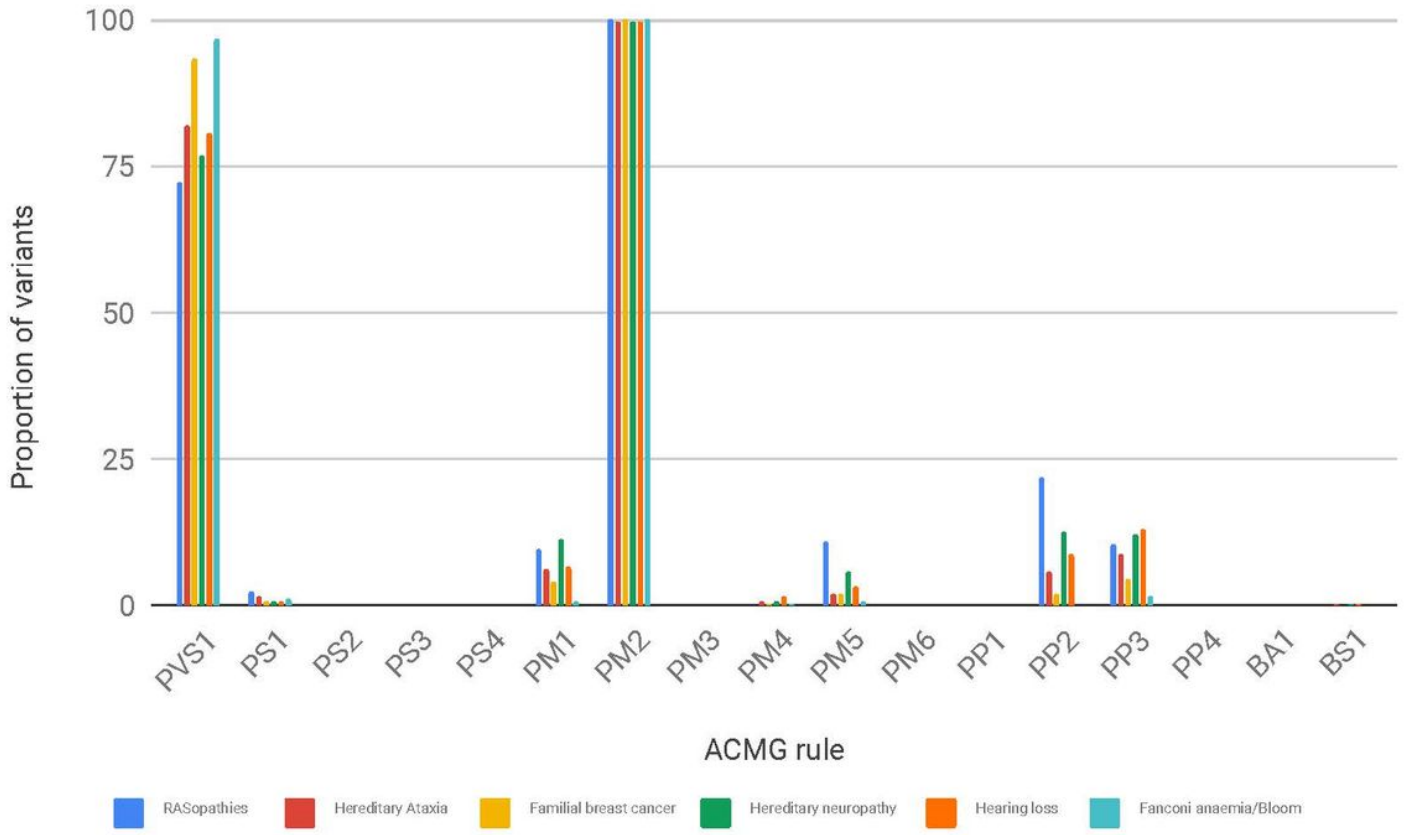

Figure 5

Distribution of ClinVar P/LP variants by aiVCE application of ACMG rules across diseases – Full dataset. ACMG, American College of Medical Genetics and Genomics; aiVCE, artificial intelligence-based Variant Classification Engine; LP, likely pathogenic; P, pathogenic

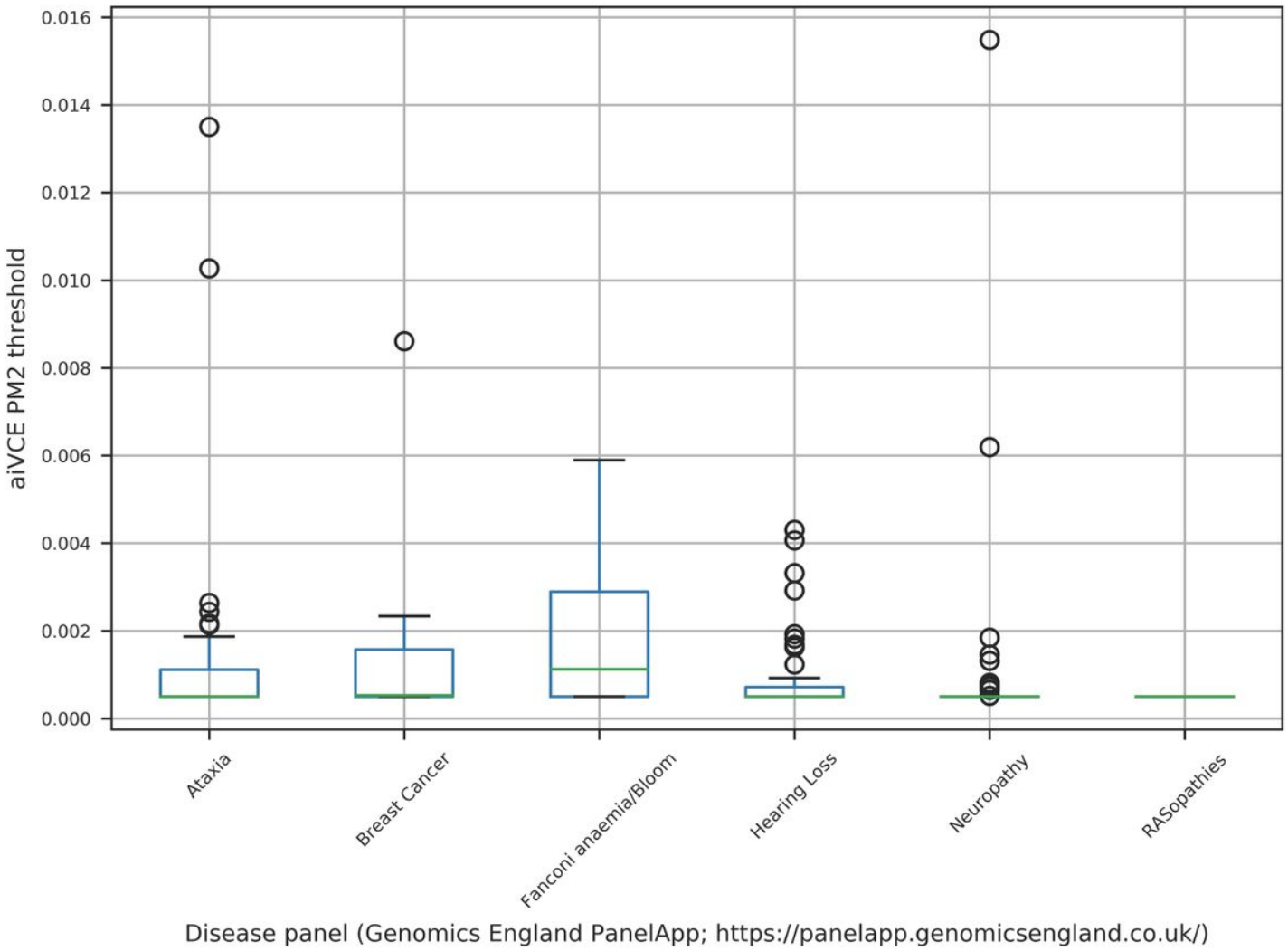

Figure 6

Boxplot: ACMG PM2 rule thresholds applied by the aiVCE across diseases - Full dataset. Only variants of genes in each panel assessed. Open circles represent single gene threshold, dark horizontal line = mean, light horizontal line = median, box = 95% confidence interval ACMG, American College of Medical Genetics and Genomics; aiVCE, artificial intelligence-based Variant Classification Engine

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplement1.xlsx