Reinterpretation of Sequence Variants Using Artificial Intelligence – Results of 2 Benchmarking Experiments

Yaron Einhorn, MSc; Adaia Kamshov, MD, MSc; Nurit Paz-Yaacov, PhD; Moshe Einhorn, MSc; Steven Harrison, PhD; Yuval Yaron, MD

1Genoox, Tel Aviv, Israel; 2Broad Institute of MIT/Harvard, Cambridge, MA, USA; 3Tel Aviv Sorsky Medical Center, Tel Aviv, Israel

ABSTRACT

Introduction: Reinterpretation of sequence variants is an important task of diagnostic laboratories, because variant classification may change over time as new evidence is accrued. Recent publications estimate the time required for reinterpretation at 1-2 hours per variant, thus creating a daunting bottleneck (Harrison S, et al., Human Mutat 2018;39:1641-9). Some variants, are exceptionally challenging to classify, owing to contradicting evidence of their pathogenicity. To address challenges in variant interpretation, a novel Artificial Intelligence (AI) Variant Classification Engine (aiVCE) was developed, based on the American College of Medical Genetics and Genomics (ACMG) Standards and Guidelines for the interpretation of sequence variants. The aiVCE is motivated by the work of recent ClinGen expert groups, which specify assertion criteria based on the unique features of particular genes or domains of interest for each variant. In addition, the aiVCE is able to resolve contradicting classifications submitted to ClinVar.

Methods: The utility of this aiVCE was evaluated against two recently published datasets of variants submitted to ClinVar (Harrison S, et al., Human Mutat 2018;39:1641-9). The first dataset included 21,199 variants with concordant classification among at least 2 of 41 selected laboratories (Concordant Dataset). Because the aiVCE relies on ClinVar consensus, PP5/BP6 rules (“reputable source recently reports variant as pathogenic/benign”) were not taken into account. The second dataset included 127 variants with outlier classifications of medically significant differences (MSDs) which were ultimately resolved following reassessment prompted by that study (Discordant Dataset). In both assessments, variants were grouped as “actionable” (pathogenic/likely pathogenic) and “non-actionable” (variant of uncertain significance/benign/benign) (VUS+LB+B).

Results: In the Concordant Dataset, 19,816/21,199 (93.5%) variants were classified by the aiVCE in agreement with ClinVar when comparing “actionable” vs “non-actionable” classifications (Table 1). Of the variants that were not in agreement, no critical discrepancies in classification were detected, e.g., no P or L variants were classified as B or L by the aiVCE. In the Discordant Dataset, when assessing the 127 variants with MSDs that were resolved after reassessment, 94 (77.2%) were classified by the aiVCE in concordance with the published post-reassessment classification (Table 2). All 33 remaining variants were classified by the aiVCE in concordance with the outlier classification. Of these, two variants (CHEK2_NM_007194.3.c:4707>C, HFE_NM_000410.3.c:845G>A) were detected by the aiVCE as “Risk susceptibility variants,” therefore, despite of their high frequency, all B frequency rules (BS1, BS2, BA1) were disabled automatically. The majority of variants classified as P/LP in ClinVar and VUS by the aiVCE were due to ACMG criteria that cannot be automatically applied, study-specific related to case-level data. Conclusions: These benchmarking results demonstrate the aiVCE’s ability to classify variants at a high rate of concordance with expert curations. The high level of agreement with challenging variants, such as those in the Discordant Dataset, makes the aiVCE ideal to expedite reinterpretation of variants accumulated over time. The aiVCE can be designed to monitor and alert the lab of variants with significant classification changes at predetermined time intervals.

Aims

- Benchmark the utility of the aiVCE as a reinterpretation tool by performing a retrospective study using challenging variants for which there were originally contradicting interpretations.
- Assess aiVCE robustness by benchmarking against large number of variants for which there is consensus between all submitters in ClinVar.

Methods

Benchmarking Experiments

Employed two recently published datasets of variants submitted to ClinVar:

- Concordant Dataset: 21,199 variants with concordant classification among ≥2 of the 41 laboratories
- Discordant Dataset: 127 variants with outlier classifications of medically significant differences (MSDs) where ≥3 laboratories classified the variant as discordant.

- Dataset #1: “Concordant” = 21,199 variants with concordant classification among ≥2 of the 41 laboratories
- Dataset #2: “Discordant” = 127 variants with outlier classifications of medically significant differences (MSDs) where ≥3 laboratories classified the variant as discordant.

- PP5/BP6 rules (“reputable source recently reports variant as pathogenic/benign”) not taken into account.
- Considered challenging variants, as there were MSDs in their interpretation.
- Ultimately resolved following laboratory reassessment.
- Three major classification levels (P/LP, VUS, and LB/B) were employed: classifications for each variant were compared to determine concordance or discordance.
- Grouped as “actionable” (P+L) vs. “non-actionable” (VUS+LB+B).

CONCLUSIONS

- Variant classification can diverge among laboratories and evolve over time, mainly due to new/additional evidence associated with the case. As such, variant reinterpretation is of great importance.
- Benchmarking results demonstrate the aiVCE’s ability to classify variants at a high rate of concordance with expert curations.
- 100% of the discordant variants were classified consistently with ≥1 of the laboratories.
- AI-based tools can greatly reduce the turnaround-time associated with classifying unambiguous variants and thus allow more time for the manual assessment required for challenging cases.
- The aiVCE can be designed to monitor and alert laboratories of variants with significant classification changes at predetermined time intervals.

References