# Validation of internal laboratory variant classifications using ClinVar

John Castiblanco, Megan Judkins, Sophie Candille, Chris Beaumont, Eric Olson, Eric A. Evans, Imran S. Haque, H. Peter Kang, Rebecca Mar-Heyming

*South San Francisco, California*

The Counsyl clinical laboratory has developed a robust variant classification protocol involving both manual and automated curation for novel variants discovered by next-gen sequencing.

## Introduction

Variant classification and data curation are evolving into a refined and elaborate field supported by a surplus of new high-throughput next generation sequencing information. New data and variants generate interpretation challenges which call for a continuous review of the recommended standards and guidelines, generally compiled and presented by the American College of Medical Genetics and Genomics (ACMG), the Association for Molecular Pathology (AMP) and the College of American Pathologists (CAP). As part of ongoing quality assurance, Counsyl sought to compare the interpretation of inherited cancer variants observed at Counsyl with their corresponding classification submitted by other diagnostic laboratories to ClinVar.

## Methods

We analyzed the classification concordance of variants observed by Counsyl with the correspondent ClinVar submissions from 14 diagnostic laboratories. Variants were separated into 2 general categories according to medical management implications (i.e., Pathogenic and Uncertain/Benign) (Figure 1).

Classification concordance was assessed by two different approaches (Figure 1):

**Approach 1:** Concordance between Counsyl observed variants and ClinVar, this analysis included submissions with at least one classification submitted (n=2,068) (i.e., Counsyl vs. ClinVar). Observed average number of submissions per variant was 2.50.

**Approach 2:** Concordance between laboratories with at least two submissions in ClinVar for the set of Counsyl observed variants (n=1,421) (i.e., Within ClinVar classification concordance). Observed average number of submissions per variant was 3.19.

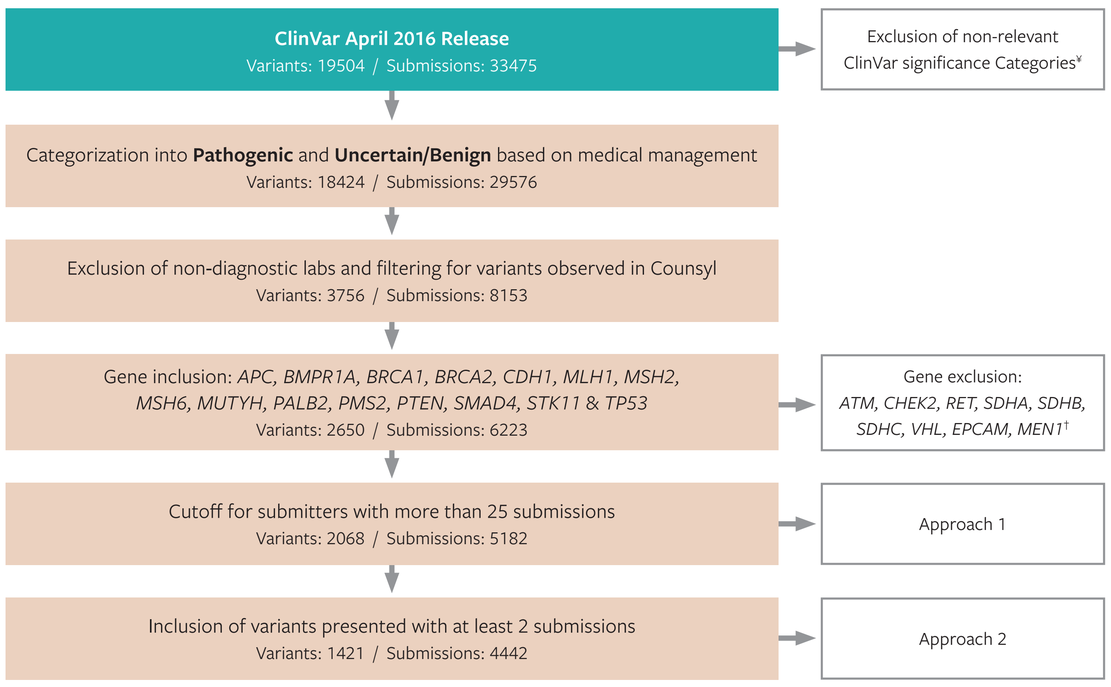


Figure 1 **Workflow showing specific steps taken for the concordance analysis**

¥ Drug response, association, risk factor, protective, affects, other and not provided clinical significance values used in ClinVar.
† Gene exclusion due to complex phenotype and/or incomplete/low penetrance. | ClinVar citation - Landrum et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016 4; 44 (D1):D862-8

## Results

**Classification concordance** was calculated as the percentage of concordant classifications among all possible pairwise comparisons per each variant. The average concordance per variant with Counsyl (i.e., Approach 1) and within ClinVar labs (i.e., Approach 2) for variants classified and submitted to ClinVar was high: 99.39% and 98.99%, respectively.

**Twenty-eight out of 2068 variants were discordant** when ClinVar submissions were compared with Counsyl classifications. Twenty-four of the 28 Counsyl discordant variants (Approach 1) were also discordant among other ClinVar submitters (Approach 2) (Figure 2).

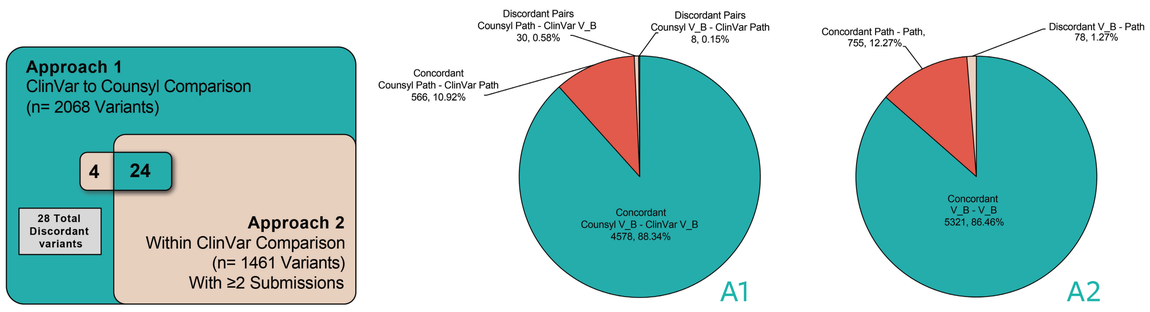


Figure 2 **Shared variant diagram and pairwise categories. Approach 1 (A1) and Approach (A2).**

**Pairwise classification categories** Concordant VUS/Benign pairs were most prevalent among pairwise comparisons for both approaches (Figure 2: A1 and A2). Moreover for the discordant pairs categories, 0.58% ($n_{Variants}$= 23) were classified as Pathogenic by Counsyl and as VUS/Benign in Clinvar and 0.15% ($n_{Variants}$= 5) as VUS/Benign at Counsyl and Pathogenic in ClinVar (Figure 2B).

Figure 3 shows the **percentage per gene for variants** included in the concordance calculations for approach 1 and 2. No significant enrichment was observed for any of the genes.

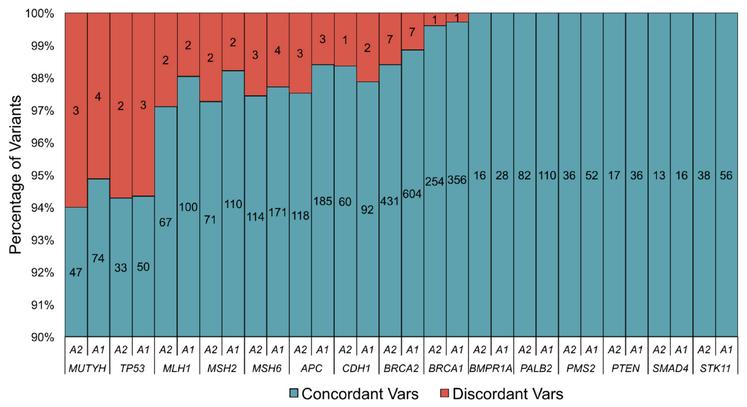


Figure 3
**Histogram depicting the number of variants included per gene in the concordance calculations for both approaches. Each bar is presented with the number of variants per category of concordance. A1: Approach 1; A2: Approach 2**

**Variant type distribution** for the compared variants showed enrichment of missense (P-value < 0.01) and depletion of silent (P-value < 0.001) type variants, for both approaches (Table 1).

| Variant type | Path % | APPROACH 1 | | APPROACH 2 | |
|---|---|---|---|---|---|
| | | Concord (%) | Discord (%) | Concord (%) | Discord (%) |
| **Missense** | 3.07 | 1117 (98.07) | 22 (1.93)* | 802 (97.69) | 19 (2.31)* |
| **Regulatory** | 0 | 35 (97.22) | 1 (2.78) | 18 (100) | 0 (0) |
| **Nonsense/Frameshift** | 97.52 | 160 (99.38) | 1 (0.62) | 121 (99.18) | 1 (0.82) |
| **Silent** | 0.21 | 477 (100)** | 0 (0) | 335 (100)** | 0 (0) |
| **Splice** | 6.27 | 251 (98.43) | 4 (1.57) | 121 (96.8) | 4 (3.2) |

Table 1 **Concordance by variant type for both approaches. Pathogenic percentage presented relative to Counsyl classification. Enrichment performed by 2x2 tables and two-tailed Fisher's exact test. p-values: *0.01 and **0.001**

## Study limitations

- Lack of regular updates and review of the publicly available variant interpretation and classification will affect concordance calculations.
- The submission and availability of only a subset of variants classified by laboratories most likely will create ascertainment bias.

## Conclusions

- The classification concordance for Counsyl with ClinVar submissions is high.
- All discordances observed in Counsyl showed disagreement among ClinVar submitters. This suggests that classifications for these variants may be controversial and worth revisiting regularly as new data becomes available.
- Although Counsyl does not use ClinVar classifications for interpretation and classification decisions, these data demonstrate the utility of a public resource in providing a platform for comparing classifications among laboratories.

**View all posters and research at**
research.counsyl.com