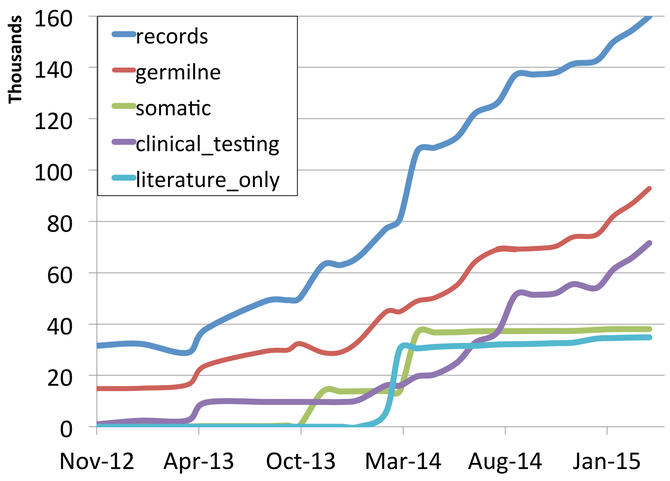
Patterns of similarity and difference among clinical interpretations from multiple laboratories in ClinVar

Shan Yang, Keith Nykamp, Yuya Kobayashi, Stephen Lincoln, and Scott Topper
Invitae, San Francisco, California, USA

Purpose

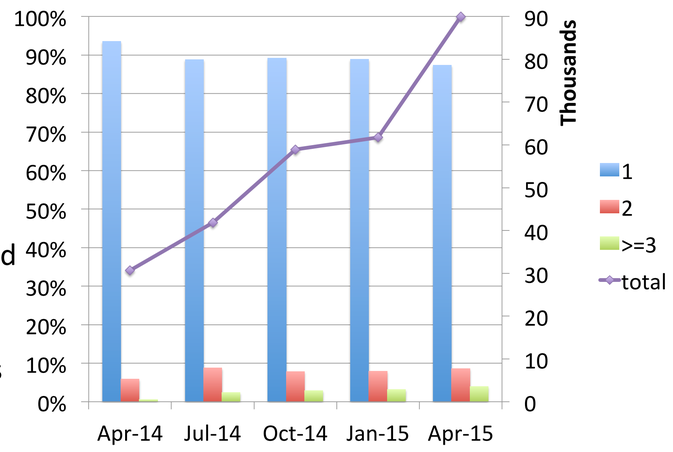
Many germline variants are quite rare and have limited published clinical and functional data available, making assessment of the clinical impact of these variants challenging. To further exacerbate this problem, after variants are observed and interpreted by a clinical lab, that information is not always accessible to the rest of the clinical and academic community. As a way to collate multiple observations of rare variants, among hundreds of independent clinical testing and academic research laboratories, the publically available ClinVar database (www.ncbi.nlm.nih.gov/clinvar/) was established. In just two-and-a-half years, the number of records submitted to ClinVar have grown approximately fivefold. To date, a large-scale comparison of the clinical assertions for variants with multiple laboratory submissions in ClinVar has not been published. In this study, we reviewed the data for a subset of ClinVar variants as a way to further understand the clinical utility of this growing and increasingly important resource.

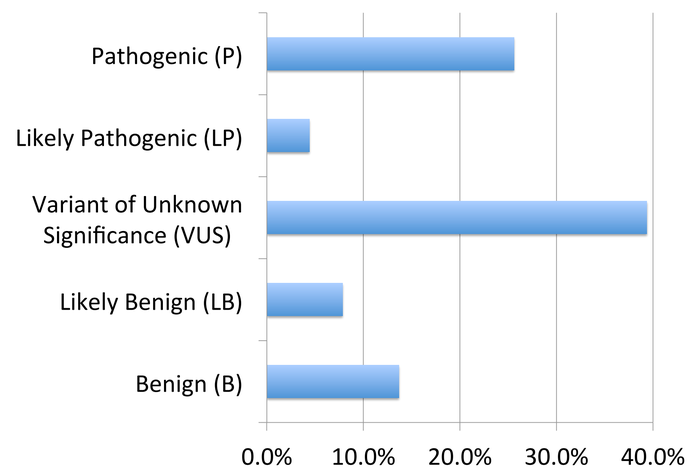
ClinVar background



Total records: From November 2012 to April 2015, ClinVar grew from 32,623 records to 159,914. The greatest growth is in records describing germline variants observed in clinical testing, and approximately 2,700 germline variant records are currently are added per month. Note that not all ClinVar records are properly tagged with their collecting methods or source.

Number of records for the same variant: Most germline variants in ClinVar have been reported by a single laboratory, excluding "literature only" submissions. Interestingly, the percentage of variants with 2 or 3+ submissions has not dramatically increased, even with the rapid acceleration of total ClinVar submissions. This suggests that the rate of new-variant discovery by clinical testing labs continues at a rapid pace and shows little sign of leveling off.





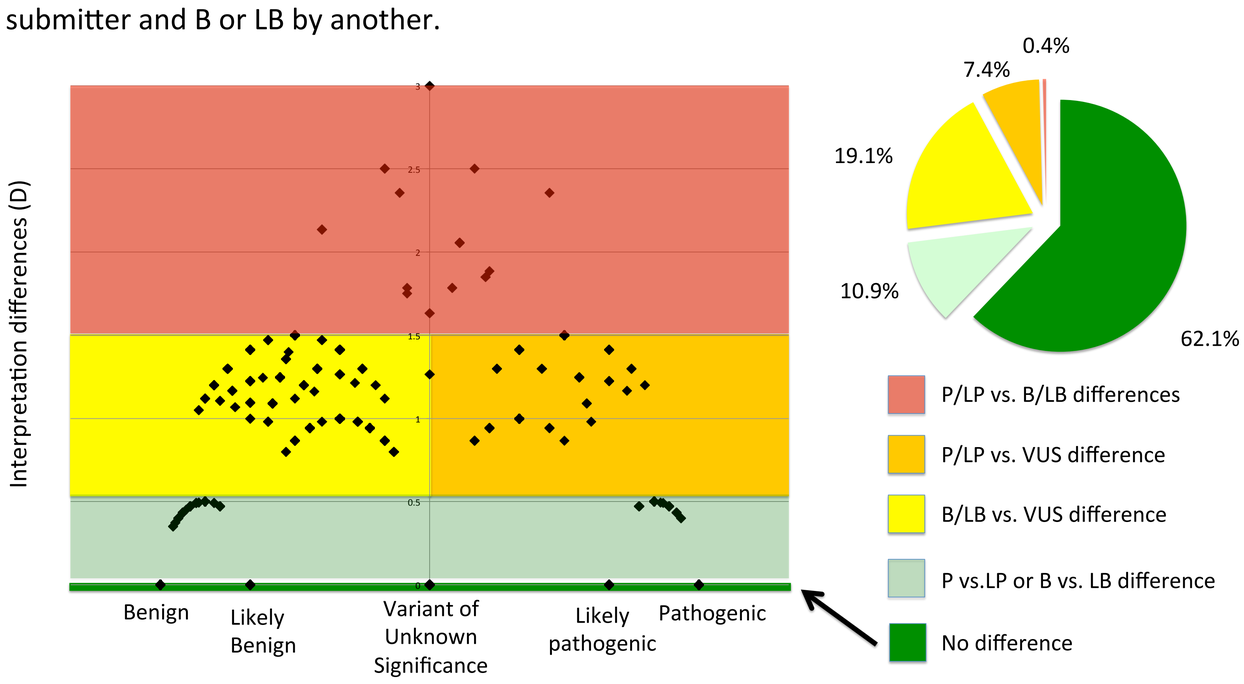
Clinical classification: Most submissions to ClinVar include a classification following the five-class schema recommended by American College of Medical Genetics (ACMG). We examined all submissions for 216 selected genes and found that a large fraction (~40%) has been classified as a variant of uncertain significance (VUS). Notably, benign variants should greatly outnumber pathogenic and VUS entries in ClinVar, but most laboratories do not submit common polymorphisms to ClinVar.

Methods

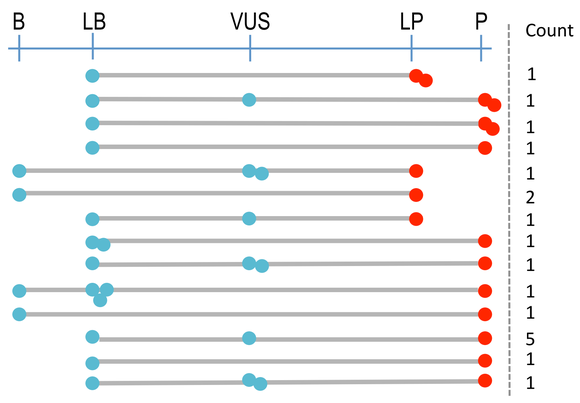
For our interpretation concordance analyses, we focused on 216 genes for which our laboratory tests and which are thus the genes we are most familiar with. These genes span hereditary cancer, cardiology, and neurology conditions. In the scatter plot, the mean score (x-axis) is calculated with Benign=-3, LB=-2, VUS=0, LP=2, and Pathogenic=3. The quantified degree of interpretation differences (D, y-axis) among the same variant is calculated in such way that D=0 represents no difference; 0<D<=0.5 represents B vs. LB or P vs. LP difference (green zone); 0.5<D<=1.5 represents B/LB or P/LP vs. VUS difference (yellow zone), and D>1.5 represents the difference being B/LB vs. P/LP (red zone). For the consensus analysis, we also filtered ClinVar submissions to those from 13 established clinical testing labs. Three or more submitters were required and consensus was defined as at least 60% of laboratories (e.g., two labs out of three, three labs out of four, four labs out of six) exactly agreeing on the five-class scale. We then conducted an "odd man out" analysis to see how often any individual submission disagreed with this consensus.

Concordance of interpretations in ClinVar

We examined the concordance of interpretations for variants submitted by two or more labs across our set of 216 genes. 62.1% of variants have no differences in interpretation among ClinVar submitters. 10.9% of variants have differences only in likelihood or strength of the assertion (e.g., B vs. LB or P vs. LP). A large percentage of variants (27.5%) are classified VUS by one submitter and either B/LB (19.1%) or P/LP (7.4%) by another. Finally, a small percentage (0.4%, n=19) includes the most extreme differences—being classified as P or LP by one submitter and B or LB by another.



Scatter plot on the left: Each point is one variant, with the x-axis being the average interpretation and the y-axis the degree of interpretation difference (see "Methods"). The green line represents no difference; the green zone is B vs. LB or P vs. LP differences. Light yellow is B/LB vs. VUS and dark yellow P/LP. vs. VUS. The red zone contains the most extreme disagreements: B/LB vs. P/LP. The pie chart shows the distribution of each degree of difference as a percentage of variants.

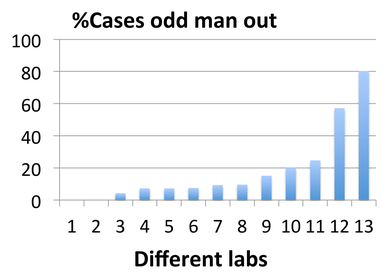


For the 19 variants (0.4%) with an extreme disagreement, we examined the degree of variability among all submitters and attempted to understand the reasons behind these differences. In some cases (4 of 19), there is a single lab that disagrees with a consensus interpretation from other labs. New evidence used in more recent submissions appears to be one cause of such differences. Mechanical errors in submissions appear to be another cause.

Most variants however showed substantial agreement. Of the 878 variants with 3+ submitters, a consensus interpretation could be derived for 818 variants. We compared the 2,714 individual submissions for these variants against this consensus and found high agreement. Most disagreements were modest (e.g., P vs LP), although again, a few were extreme (e.g., LB vs P).

Consensus interpretation

Individual ClinVar submitters	Pathogenic	LP	VUS	LB	Benign	Total	# Interpretations
Pathogenic	98.1%	1.2%	0.5%	0.2%	0.0%	100%	429
LP	28.0%	60.0%	12.0%	0.0%	0.0%	100%	25
VUS	0.4%	0.3%	93.3%	3.9%	2.2%	100%	743
LB	0.3%	0.0%	11.8%	56.2%	31.7%	100%	363
Benign	0.0%	0.0%	2.2%	3.7%	94.1%	100%	1154
						Total	2714



Certain ClinVar submitters were often in disagreement with consensus. In some cases (such as Lab 13), the trend was toward more conservative interpretations (more often VUS, and more often likely vs. strongly P or B) than consensus. In other labs, the pattern of disagreements was less systematic.

Conclusion

ClinVar is a valuable resource for clinical laboratories and has been growing rapidly as submissions continue to increase. Most variants in ClinVar have a clinical interpretation asserted by only a single lab, and a large fraction of submissions are for VUS (variants of uncertain significance). When multiple submissions are available for a variant, concordance is high, although disagreements can be quite substantial. Users of ClinVar should be aware of these limitations. These findings also emphasize the value of data sharing and peer review and should implore more laboratories to release de-identified interpreted variants to centralized open resources such as ClinVar (whenever they are legally able). Curated databases, such as those based on LOVD, play a different, but clearly important, role as well.