

Sources of discordance among germ-line variant classifications in ClinVar

Shan Yang, PhD¹, Stephen E. Lincoln, BS¹, Yuya Kobayashi, PhD¹, Keith Nykamp, PhD¹, Robert L. Nussbaum, MD, FACP, FACMG^{1,2} and Scott Topper, PhD, FACMG¹

Purpose: ClinVar is increasingly used as a resource for both genetic variant interpretation and clinical practice. However, controversies exist regarding the consistency of classifications in ClinVar, and questions remain about how best to use these data. Our study systematically examined ClinVar to identify common sources of discordance and thus inform ongoing practices.

Methods: We analyzed variants that had multiple classifications in ClinVar, excluding benign polymorphisms. Classifications were categorized by potential actionability and pathogenicity. Consensus interpretations were calculated for each variant, and the properties of the discordant outlier classifications were summarized.

Results: Our study included 74,065 classifications of 27,224 unique variants in 1,713 genes. We found that (i) concordance rates differed among clinical areas and variant types; (ii) clinical testing

methods had much higher concordance than basic literature curation and research efforts; (iii) older classifications had greater discordance than newer ones; and (iv) low-penetrance variants had particularly high discordance.

Conclusion: Recent variant classifications from clinical testing laboratories have high overall concordance in many (but not all) clinical areas. ClinVar can be a reliable resource supporting variant interpretation, quality assessment, and clinical practice when factors uncovered in this study are taken into account. Ongoing improvements to ClinVar may make it easier to use, particularly for nonexpert users.

Genet Med advance online publication 1 June 2017

Key Words: clinical genetic testing; ClinVar; concordance; data sharing; variant interpretation

INTRODUCTION

Determining the consequences of genetic variation on human health and disease is central to modern clinical molecular diagnostics. As the technical, legal, and commercial landscapes evolve, clinical molecular genetics is undergoing a renaissance: more laboratories are serving more clinicians and are testing more patients for more genes than ever before.^{1–3} However, this surge also increases the chance that laboratories may disagree about the clinical interpretation of analytic test results.

Two important public efforts have been launched to promote consistency and accuracy in variant classification. The first is the American College of Medical Genetics and Genomics (ACMG) and Association for Molecular Pathology (AMP) 2015 joint guidelines for variant classification,⁴ which aim to provide a structured methodology for variant interpretation. The second is ClinVar, the National Institutes of Health open-access database of clinically observed variants and their classifications.^{5,6} Since its release in 2012, ClinVar has grown to host hundreds of thousands of variant classifications from hundreds of academic and commercial diagnostic laboratories, consortia, literature curation efforts, and research groups.^{7,8} Such sharing of clinical genetic data is recommended by the ACMG, the American Medical

Association, the National Society of Genetic Counselors, and other organizations as a standard part of responsible clinical laboratory practice.^{9–11}

ClinVar provides a central platform for objective analyses of the interlaboratory reproducibility of variant classification, as well as a mechanism for collaborative identification and resolution of discrepancies.¹² Many recent studies have evaluated the reproducibility of variant classification using ClinVar and other data sources.^{8,13–23} Conclusions from these studies differ, and controversy remains about the true rate, nature, and clinical significance of classification differences. For example, Vail et al.¹³ reported high discordance among *BRCA1* and *BRCA2* classifications in various public databases, concluding that these resources should be “precluded from clinical use.” However, both a systematic analysis of ClinVar¹⁴ and a large prospective clinical study¹⁵ showed high concordance in these two genes. Balmaña et al.¹⁶ found high discordance in non-*BRCA1/2* cancer genes, whereas Maxwell et al.¹⁷ found high concordance, as did our own re-analysis of the data available for Balmaña’s variants.¹⁸ Van Driest et al.¹⁹ found high discordance in arrhythmia genes, although the specific data they used may have been overclassified, which is inconsistent with the ACMG/AMP guidelines.²⁰ Finally, although Amendola et al.,²¹ Garber et al.,²² and Harrison

¹Invitae, San Francisco, California, USA; ²Volunteer Clinical Faculty, University of California, San Francisco, California, USA. Correspondence: Stephen E. Lincoln (steve.lincoln@me.com)

The first two authors contributed equally to this paper.

Submitted 24 January 2017; accepted 31 March 2017; advance online publication 1 June 2017. doi:10.1038/gim.2017.60

*et al.*²³ found classification discordances in various genes, they also found that the sharing of information among laboratories led to far greater concordance even in difficult classification situations.

In this study, we carefully analyzed variant classification data in ClinVar, quantifying concordance both across and within various medical specialties in terms of potential clinical impact. When classification differences existed, we investigated factors that may have contributed to discordance, such as the source of data, date of classification, penetrance, and variant type. Our results may help design future studies of classification concordance and should inform database users of critical considerations for using these public resources.

MATERIALS AND METHODS

Our analyses used the ClinVar October 2016 XML file.²⁴ For each variant interpretation, we extracted variant identification, preferred Human Genome Variation Society variant description, gene name, variant classification, submission type (“collection method” field), classification date (“last evaluation” field), classification criteria (“review status” and “assertion method” fields), submitter name, and submitter comments (if provided), including the summary evidence description. This information is also available in the ClinVar user interface (**Figure 1**). We also examined the clinical condition field, although we used our own high-level clinical categories in these analyses (**Supplementary Data** online). We identified cases in which a variant was represented by multiple identifications and merged these records. When multiple submissions for the same variant were present from the same submitter, only the most recent was used. Somatic (as opposed to germ-line) entries, and entries without classifications, were ignored.

All variants were assigned a functional category as follows: (i) small protein sequence alterations included both missense variants and, much less frequently, in-frame indels; (ii) truncating variants included both nonsense (stop – gain) and out-of-frame indels; (iii) silent variants were those in coding regions that do not change amino acid sequence; (iv) splice-site variants were intronic changes within two bases of an exon boundary; and (v) noncoding variants (including those in introns, untranslated regions, and intergenic regions) were combined. This simplified framework was intended to generally distinguish straightforward classification scenarios from complex ones. In clinical reporting, a more sophisticated assessment must be performed.

The Exome Aggregation Consortium database^{25,26} version 0.3²⁷ provided population allele frequencies. In total, 29,286 variants had ClinVar classifications from two or more submitters. We removed 2,062 benign polymorphisms, defined as variants that had both a greater than 5% allele frequency in the Exome Aggregation Consortium database and only benign classifications in ClinVar. These prevalent, easily interpreted, and clinically irrelevant variants^{4,28–30} may be included or omitted depending on the submitter. Excluding such variants from this study improves consistency,

although we thus underestimate concordance relative to the complete spectrum of variants in patients.

Classifications are recorded in ClinVar using the ACMG/AMP five-tier system⁵ with some exceptions. Laboratory-specific categories (e.g., “deleterious” instead of “pathogenic,” or “polymorphism” instead of “benign”) are mapped to the ACMG/AMP nomenclature by submitters or by ClinVar.³¹ In our analyses, we considered pathogenic and likely pathogenic classifications equivalent. Similarly, benign and likely benign were considered equivalent. Other ClinVar classifications, including “risk factor,” “drug response,” and “other” were considered not pathogenic and not clinically actionable.

We used a two-tailed *z*-test to compute *P* values for differences in observed proportions. *P* values less than 0.0002 were not computed exactly. Ninety-five percent confidence intervals (CIs) were calculated using the Wilson method.

RESULTS

Our data set included 27,224 variants classified by two or more ClinVar submitters. These spanned 1,713 genes and totaled 74,065 classifications from 429 ClinVar submitters. Of these variants, 17,115 had classifications from exactly two submitters, and 10,109 had classifications from three or more. The complete data set is provided in the **Supplementary Data**.

Variant interpretations in ClinVar are highly concordant

We analyzed the concordance of classifications at two levels of detail: potential clinical actionability and pathogenicity. Our actionability analysis distinguished two groups: positive (pathogenic or likely pathogenic) and potentially clinically actionable versus not positive and not actionable (benign, likely benign, or variants of uncertain significance (VUS)). Our pathogenicity analysis considered three groups: positive (identical to the set defined above), uncertain (VUS), and negative (benign and likely benign). These two analyses mimic those of previous studies.^{15–20} The analyses described below are based on actionability unless otherwise specified. For both the actionability and pathogenicity comparisons, if 100% of the classifications were in the same group, the variant was said to have complete agreement. If at least two-thirds of the classifications were in the same group (i.e., 2 of 2, 2 of 3, 3 of 4, 4 of 5, 4 of 6, etc.), the variant was said to have majority consensus, potentially with outliers. Many of our analyses examined the characteristics of these outliers. Note that a set of variants with complete agreement is always a subset of those with majority consensus.

Our analysis shows that variant classifications in ClinVar are highly concordant. Considering actionability, 96.7% of variants (CI: 96.5–96.9%) reached majority consensus, and 94.1% (CI: 93.9–94.4%) had complete agreement among all submitters (**Figure 2**). Considering pathogenicity, 89.3% (CI: 88.9–89.6%) had majority consensus, and 81.0% (CI: 80.5–81.4%) had complete agreement. The lower concordance observed on a pathogenicity basis demonstrates that classification differences between VUS and benign (including likely

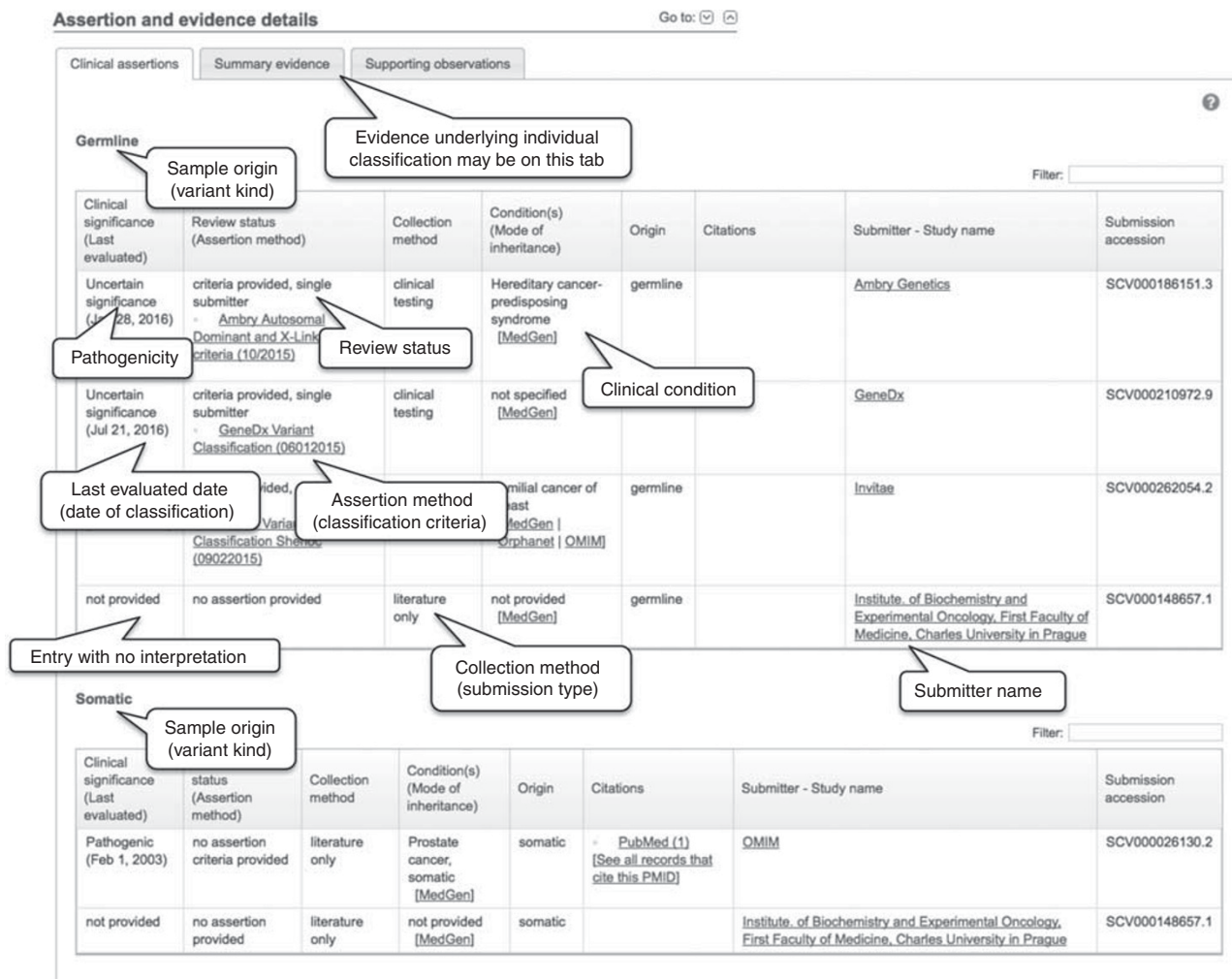


Figure 1 ClinVar variant NM_007294.3(BRCA1):c.4327C > G (p.Arg1443Gly). Key data fields used in our analyses are indicated. Where common names differ from ClinVar terminology, both are included. From <https://www.ncbi.nlm.nih.gov/clinvar>, accessed 6 March 2017.

benign) are much more common than differences between either VUS or benign and pathogenic (including likely pathogenic). In both pathogenicity and actionability analyses, interpretation differences and certain reporting differences may cause discordance—our study methodology cannot distinguish between these cases (as one example, see *GALT*:c.940A > G, discussed by Garber *et al.*²²).

Nonclinical submissions account for many of the discordances

Most ClinVar classifications are of one of four types: clinical testing (79.7%), literature only (8.8%), curation (5.8%), or research (5.1%). These categories are self-reported by each submitter and are usually (but not always) used consistently. Most clinical laboratories submit clinical testing classifications, although some submit variants tagged as literature only or curation. Reference databases, including the Online Mendelian Inheritance in Man (OMIM) database,³² are usually tagged as literature only. Consortia submissions vary: for example, ENIGMA (Evidence-based Network for the

Interpretation of Germline Mutant Alleles)³³ submits curation records, whereas InSiGHT (the International Society for Gastrointestinal Hereditary Tumours)³⁴ submits research records. ClinVar users should be aware that classifications of types other than clinical testing may not follow the rigorous criteria used by most diagnostic laboratories.

We collected variants that achieved majority consensus with one or more outliers (i.e., that were not in complete agreement) and examined these outliers in detail. There were 702 such variants with 739 outlier classifications. Although literature-only submissions composed merely 8.8% of our data set, they made up 31.4% of the outliers. Overall, 3.6% of literature-only submissions were outliers—six times the 0.6% rate of clinical testing submissions (*P* < 0.0002). Research and curation submissions were also disproportionately represented among outliers (Figure 3a).

ClinVar contains many genes not commonly tested in current clinical practice. To confirm our results in well-studied, commonly tested genes, we repeated this analysis using only the 23 genes (*BRCA1*, *BRCA2*, and others; see

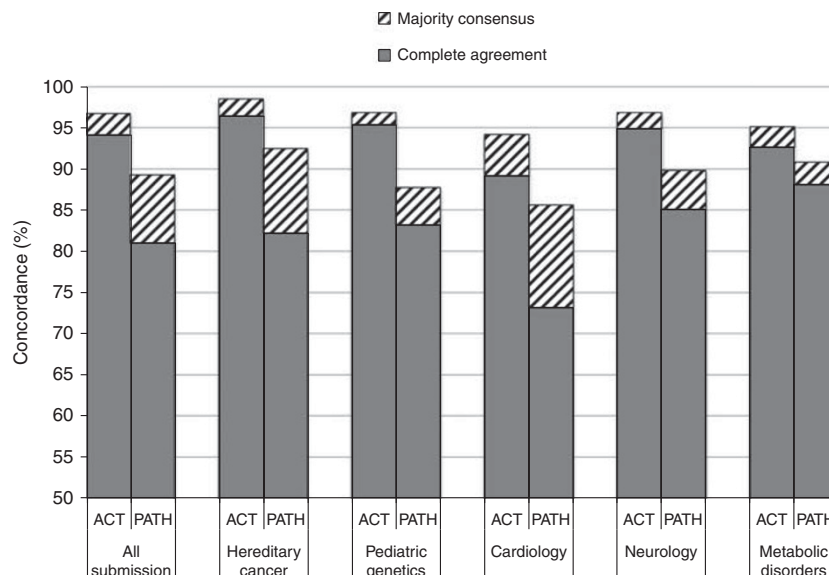


Figure 2 Concordance by clinical area. Variant classification concordance measured as a fraction of variants for all genes (leftmost pair) and for genes in each clinical area (other pairs). The left bar in each pair (labeled ACT) represents our actionability analysis, and the right bar (PATH) indicates our pathogenicity analysis (see text for details). Note that the y axis starts at 50%.

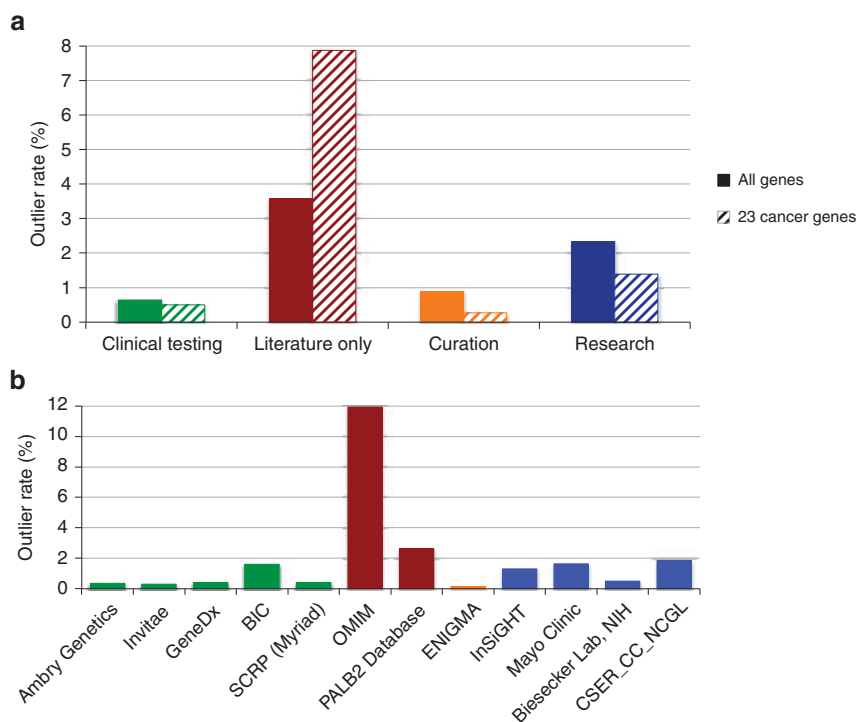


Figure 3 Outlier rate by submitter. (a) Outlier rate (fraction of all classifications discordant with the majority consensus) calculated on an actionability basis for submissions of each major type. The left bar in each pair represents all ClinVar variants, and the right bar is restricted to the 23 hereditary cancer genes listed in the National Comprehensive Cancer Network guidelines (see text and **Supplementary Table S1**). (b) Outlier rate for specific submitters with the largest number of classifications in the 23 cancer genes. In the current data set, these clinical testing submitters each contributed more than 1,000 variants; submitters of other types provided more than 100 variants. More than 87% of all clinical testing submissions and more than 90% of all literature-only, curation, and research submissions are represented by the submitters in (b). BIC, Breast Cancer Information Core; ENIGMA, Enhancing NeuroImaging Genetics through Meta-Analysis; NIH, National Institutes of Health; OMIM, Online Mendelian Inheritance in Man; SCRIP, Sharing Clinical Reports Project.

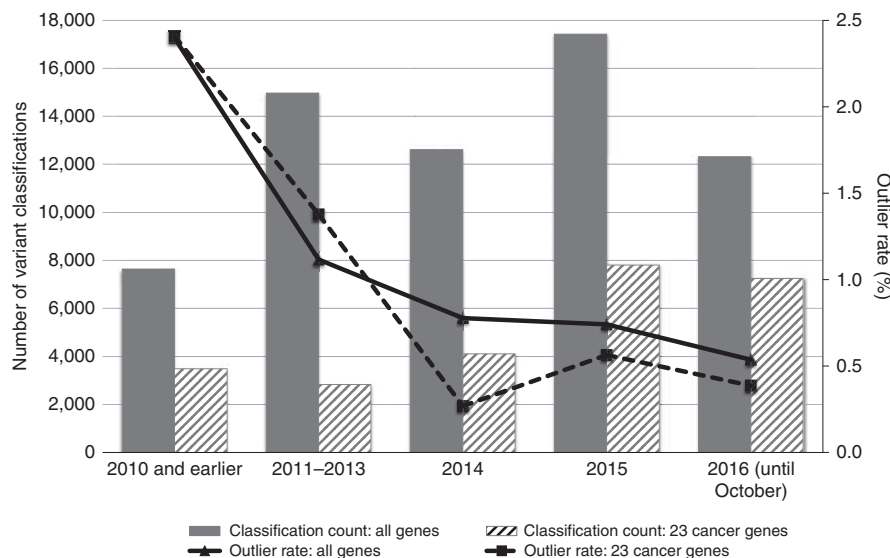


Figure 4 Outlier rate by classification date. Lines show the outlier rate for variants in each date range. Bars show the number of variant classifications for each range. Solid bars and lines show the complete ClinVar data set, whereas the hatched bars and dotted lines show only data for the 23 hereditary cancer genes listed in the National Comprehensive Cancer Network guidelines (see text and **Supplementary Table S1**).

Supplementary Table S1 included in management guidelines for hereditary breast, ovarian, colorectal, and related cancers. Variants in these 23 genes ($n = 8,765$; 32% of the total) had particularly high rates of majority consensus (98.8%) and complete agreement (96.9%). Literature-only entries accounted for 17.2% of the outliers despite being only 1.4% of the submissions. More than 7.9% of literature-only submissions were outliers compared with 0.5% of clinical testing submissions, a 16-fold difference ($P < 0.0002$; **Figure 3a**). Discordance was low among high-volume clinical laboratory submissions, with literature-only OMIM entries often discordant (**Figure 3b**).

Older classifications are often discordant

Variant classifications are made at a specific point in time; however, scientific evidence accrues continually. Most classifications in our analysis (88.2%) dated from 2011 or later, although some were much older (as early as 1976). Five years (2011–2016) is a long time in the field of genetics, not to mention 40 years (1976–2016).

We created date ranges with roughly similar numbers of classifications and compared the outlier rates between ranges. As above, outliers were computed through comparison with all data in the set, including newer submissions. Old classifications were often outliers, including 2.4% of pre-2011 and 1.2% of 2011–2013 submissions, compared with 0.8% of 2014 and 0.5% of 2016 submissions—a significant decline (P values were, at most, 0.0032 between pairs of ranges; **Figure 4**).

Combined effect of interpretation source and age

Because major sources of discordance included both non-clinical submissions and older data, we repeated our overall

concordance analysis considering only recent classifications from clinical testing. There were 20,160 variants with multiple classifications from clinical testing alone, and 57% of these (11,535) had multiple submissions from 2014 or later. Compared with ClinVar overall, concordance was significantly higher in these data, with 98.4% (CI: 98.2–98.6%) of variants having majority consensus and 97.2% (CI: 96.8–97.5%) reaching complete agreement (**Figure 5**).

We examined whether submissions with interpretation criteria (“assertion methods” in ClinVar terminology) were more consistent than those without. These were indeed highly concordant (only 0.6% outliers), although this field is often populated for recent, clinical submissions (and less often others), which explains the similarity between this result and that shown in **Figure 4**. The outlier rate was not significantly different between recent (2015–2016) clinical submissions with and without interpretation criteria ($P > 0.05$).

Concordance rates differ substantially by clinical area

The issues faced in variant classification differ among clinical areas. We assigned genes to one of five specialties based on the primary disease caused by pathogenic variants in that gene: hereditary cancer, cardiology, neurology, metabolic disorders, and unrelated pediatric conditions (**Supplementary Data**). Hereditary cancer genes had the highest majority consensus rate (98.5%), and genes related to cardiology and metabolic disorders had the lowest rates (94.2 and 95.1%, respectively; **Figure 2**).

The greatest differences in majority consensus rate between our actionability and pathogenicity analyses were observed in pediatric disease (a decrease of 9.1%) and cardiology (a decrease of 8.6%). The greatest differences in complete agreement rate between actionability and pathogenicity were

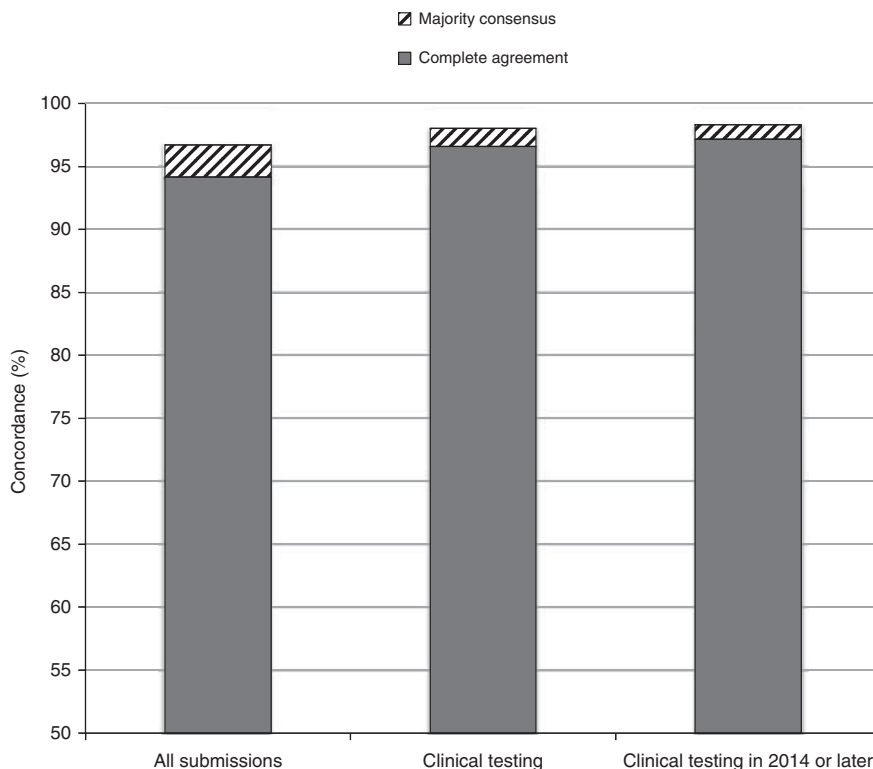


Figure 5 Concordance for ClinVar and subsets. Variant classification concordance measured as a fraction of variants for all of ClinVar and for subsets of ClinVar filtered by submission type and classification date. Concordance is calculated on an actionability basis (see text).

seen in cardiology and cancer (decreases of 14.3 and 12.2%, respectively; **Figure 2**). Collectively, these results demonstrate variability in classification at the border between VUS and benign or likely benign within these specialties.

In contrast to the 23-gene hereditary cancer subset described above, a subset of 25 commonly tested cardiology genes (**Supplementary Table S2**; 2,037 variants; 7.5% of the total) was analyzed and showed far lower concordance than the overall rates in ClinVar: only 90.9% majority consensus and 82.8% complete agreement. Outlier rates were relatively high for all submission types in these genes: clinical testing (2.3%), curation (3.0%), literature only (4.1%), and research (7.0%).

Discordance varies by variant type

The difficulty of classification varies substantially by variant type. Certain classes (e.g., truncating variants) are often simple to assess and others (e.g., missense variants) are more difficult. In our analysis, silent changes had the highest concordance (99.8%) and were almost always classified as benign. Truncating variants, which were almost always pathogenic, had the second highest concordance (98.8%; **Supplementary Figure S1**). Protein sequence changes (mostly missense variants) and splice-site variants had the lowest concordance rates (94.7 and 97.5%, respectively).

ClinVar submitters have different policies regarding which variants they contribute, and deliberate or unintentional

biases are certainly possible. If a submitter were to provide only unambiguous classifications, its concordance rate would be artificially increased. We examined the distribution of variants, looking for gross signs of bias among submission types. More than half (54.0%) of the variants reported from clinical testing were missense (**Supplementary Figure S2**), a challenging and low-concordance type. Although missense variants were a slightly larger fraction (69.7%) of the literature-only submissions, this difference does not explain the dramatically higher discordance rate in the literature-only data. Indeed, this higher fraction may be a simple function of the underreporting of silent variants within literature-only submissions. Silent variants do not substantially inflate concordance in clinical testing data, because even if they are omitted, the majority consensus rate in these submissions (alone) would decrease only slightly from 98.0 to 97.4%.

Low-penetrance variants have high discordance rates

Some variants are challenging to classify, not because the evidence is disputed, but because none of the ACMG/AMP categories accurately describes the variants' effects. Notable examples include low-penetrance pathogenic variants, which confer a relatively modest risk of disease and do not have specific criteria in the 2015 guidelines.⁴ Some laboratories classify such variants as pathogenic and add notes describing penetrance, although such notation is not always submitted to ClinVar. Other laboratories classify these variants as "risk

factors,” although this category is more often used for exceptionally low-penetrance associations (e.g., those uncovered by genome-wide association studies). Still others report these variants as VUS or likely benign.

We identified such variants by searching summary evidence descriptions for the words “low penetrance,” or identifying variants for which at least one submitter used the risk factor classification. Although this approach did not identify all low-penetrance variants, it provided a sufficient data set for analysis. We identified 238 such variants with 803 total classifications. Only 78.2% of these variants reached majority consensus for actionability, and only 49.2% reached consensus for pathogenicity. Compared with the overall rates of 96.7 and 89.3%, respectively, these variants contributed a significantly disproportionate fraction of disagreements ($P < 0.0002$ in both comparisons).

DISCUSSION

Since its release in 2012, ClinVar has grown rapidly into a key resource for clinical genetics.^{5,6,8,12} However, its use in assessing the consistency of variant classification has been controversial. Some studies of ClinVar report high discordance,^{13,16,19} although others report high concordance.^{8,14,17} Our analysis shows that classification differences in ClinVar are often explained by readily observable properties of the data.

First, not all public data deserve equal credence, an observation that may seem obvious. In our study, classifications simply excerpted from published literature or imported from research efforts were frequently discordant with formal classifications produced by clinical laboratories. In particular, variant interpretations from OMIM³², one of the top sources of data in ClinVar, accounted for a substantial fraction of disagreements. OMIM is a valuable resource for information about Mendelian disease genes and their associated conditions and molecular mechanisms. Although specific variants are often also described, OMIM makes no claim to classify these variants according to clinical standards. Unfortunately, this aspect of OMIM classifications in ClinVar can easily be misunderstood. OMIM also includes a substantial quantity of old data.

Second, variant interpretations undergo constant refinement as clinical and experimental evidence accumulates. Compared with recent classifications, older classifications were, unsurprisingly, much more likely to be in conflict in our study. It is recommended that laboratories both update clinicians when variants are reclassified and submit reclassifications to databases such as ClinVar. However, database users must be aware that entries can become outdated, and further investigation may be required to determine if and when this has occurred. Indeed, an important use of ClinVar is to identify new classifications that differ from previous ones (from the same or a different source) possibly because of new scientific evidence.

Third, although we found high agreement across clinical areas, the rates varied significantly (e.g., concordance in

cancer genes was much higher than that in cardiology genes). Although the reasons for this variability remain unclear, they may include differences in the nature and complexity of molecular mechanisms and clinical phenotypes, the quantity of available case reports and functional literature, and the length of time for which genetic testing has been part of common clinical practice, among others. A detailed investigation of such sources of discordance was beyond the scope of our study, but would certainly be valuable in larger sets of variants from ClinVar, perhaps following the models of Amendola *et al.*,²¹ Garber *et al.*,²² and Harrison *et al.*²³

A fourth factor contributing to discordance was the difference in both terminology and classification criteria applied to low-penetrance alleles. This finding is consistent with the observation by Balmaña *et al.*^{16,18} that low-penetrance variants account for a substantial fraction of discordances, even though the impact of such differences on clinical management is relatively limited when current guidelines are followed.

Finally, the spectrum of variant types in each gene was a factor. Rare missense variants are both numerous and collectively prevalent^{8,14,25} and can be among the most challenging variants to classify. For diseases in which loss of function is the key mechanism, only a small fraction of rare missense variants are pathogenic, possibly improving concordance on an actionability basis, although many are classified as (nonactionable) VUS owing to limited evidence against pathogenicity. This challenge may reduce concordance on a pathogenicity basis.

Small changes to ClinVar would help users distinguish variants for which clinical consensus exists from variants for which meaningful disagreements may be present. Strongly distinguishing between clinical and literature/curation/research submissions (and indeed further standardizing the use of these tags among submitters) would help guide users toward the most trustworthy submissions. Clear indications that a particular classification may be out of date would be helpful, as would standardized terminology for pathogenic but low-penetrance variants (until revised guidelines on this topic are available). Our two comparison types—actionability and pathogenicity—were each useful in different ways, and a ClinVar flag indicating actionability differences would be helpful in addition to the existing “conflicting interpretations” flag (which mimics our pathogenicity analysis). The ability to display conflicts incorporating the factors described in this study would certainly be helpful, although such a feature may not be simple to implement.

The presence of an assertion method in ClinVar, used by submitters to describe their general classification methodology, sheds little light on concordance beyond the factors described above. Although certainly useful, the value of this field as a simple quality indicator may be diminishing, as more submitters of various types fill it in. Far more valuable, and far less frequently populated, is the summary evidence description field intended to provide a specific rationale for

each individual classification. Per ACMG/AMP guidelines⁴ (specifically, PP5 and BP6), assertions without these details can be considered only “supporting” (i.e., weak) evidence. We suggest that this data field be made both more prominent and easier to access by moving it onto the main variant information page. Perhaps the stars awarded to any submission should take into account whether these details have been provided. At present, few ClinVar submitters populate this field (of the largest clinical submitters, only GeneDx, Harvard’s Laboratory for Molecular Medicine, and Invitae routinely do so), but these ClinVar changes could inspire others.

Despite the caveats, we found a high degree of concordance among variant classifications in ClinVar. Concordance was even higher when the factors described herein were taken into account. Concordance was particularly high for certain commonly tested genes, such as those underlying hereditary cancer syndromes. This consistency is remarkable, particularly given the substantially lower consensus rates reported for other types of medical interpretation (e.g., interpretations of breast biopsies or mammograms).^{35,36} Previous studies reporting high levels of discordance^{13,16,19} among variant classifications largely ignored the factors uncovered in our analysis and thus greatly underestimate the extent to which consistent, high-quality classifications are being produced and contributed by many diagnostic laboratories.

These studies can also mischaracterize the way ClinVar is used in practice; for instance, promulgating the false view that variant classifications are often simply copied from ClinVar into clinical reports.^{13,37} In our experience, laboratory directors are well aware that public databases must only be used appropriately. Moreover, many (though not all) laboratories consider these databases crucial to achieving consensus regarding the pathogenicity of the millions of DNA variants being identified in the hundreds of thousands of patients undergoing genetic testing each year.¹² We therefore find it unfortunate that certain laboratories with proprietary databases not only fail to contribute data,^{38,39} contrary to recommendations,^{9–11} but also enact specific policies prohibiting ordering clinicians from doing so.⁴⁰

The lack of complete consensus in ClinVar should not be exaggerated or allowed to obfuscate the obvious: free and open access to de-identified clinical testing data provides the clinical community with an invaluable and unique tool with which to share evidence, make and revise judgments, reach consensus, and ultimately provide the best medical care to patients undergoing genetic testing.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

ACKNOWLEDGMENTS

We are deeply grateful to the laboratories and other groups who are working to advance medical care by submitting data to the

ClinVar database. We thank Melissa Landrum and many other members of the ClinVar staff and ClinGen project for answering our technical questions and for their hard work on this important resource. We thank Nancy Jacoby, John Garcia, and Tim Chiu (of Invitae) for assistance with this study.

DISCLOSURE

All authors are employees of Invitae, a laboratory offering clinical genetic testing services. This project was funded by Invitae.

REFERENCES

1. National Institutes of Health Genetic Testing Registry. <http://www.ncbi.nlm.nih.gov/gtr/>. Accessed 12 January 2017.
2. Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* 2013;14:295–300.
3. Rehm HL, Hynes E, Funke BH. The changing landscape of molecular diagnostic testing: implications for academic medical centers. *J Pers Med* 2016;6:8.
4. Richards S, Aziz N, Bale 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;17:405–424.
5. Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* 2016;44:D862–D868.
6. Harrison SM, Riggs ER, Maglott DR. Using ClinVar as a resource to support variant interpretation. *Curr Protoc Hum Genet* 2016;89:8.16.1–8.16.23.
7. ClinVar. ClinVar Submissions. <https://www.ncbi.nlm.nih.gov/clinvar/submitters/>. Accessed 12 January 2017.
8. Yang S, Cline M, Zhang C, Paten B, Lincoln SE. Data sharing and reproducible clinical genetic testing: successes and challenges. *Pac Symp Biocomput* 2016;22:166–176.
9. American Medical Association. Proceedings of the 2013 Annual Meeting of the House of Delegates. Resolution 519 – Genome Analysis and Variant Identification. 2013:612. https://www.ama-assn.org/sites/default/files/media-browser/public/hod/a13-reference-committee-reports_0.pdf. Accessed 12 January 2017.
10. National Society of Genetic Counselors Position Statements: Clinical Data Sharing, 2015. <http://nsgc.org/p/bl/et/blogid=47&blogaid=330>. Accessed 12 January 2017.
11. ACMG Board of Directors Laboratory and clinical genomic data sharing is crucial to improving genetic health care: a position statement of the American College of Medical Genetics and Genomics. *Genet Med*; e-pub ahead of print 5 January 2017.
12. Rehm HL, Berg JS, Brooks LD, et al. ClinGen—the clinical genome resource. *N Engl J Med* 2015;372:2235–2242.
13. Vail PJ, Morris B, van Kan A, et al. Comparison of locus-specific databases for BRCA1 and BRCA2 variants reveals disparity in variant classification within and among databases. *J Community Genet* 2015;6:351–359.
14. Lincoln SE, Yang S, Cline MS, et al. Consistency of BRCA1 and BRCA2 variant classifications among clinical diagnostic laboratories. *JCO Precis Oncol* 2017;1:1–10.
15. Lincoln SE, Kobayashi Y, Anderson MJ, et al. A systematic comparison of traditional and multigene panel testing for hereditary breast and ovarian cancer genes in more than 1000 patients. *J Mol Diagn* 2015;17:533–544.
16. Balmaña J, Digiovanni L, Gaddam P, 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;34:4071–4078.
17. Maxwell KN, Hart SN, Vijai J, et al. Evaluation of ACMG-guideline-based variant classification of cancer susceptibility and non-cancer-associated genes in families affected by breast cancer. *Am J Hum Genet* 2016;98:801–817.
18. Nussbaum RL, Yang S, Lincoln S. 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;35:1259–1261.
19. Van Driest SL, Wells QS, Stallings S, et al. Association of arrhythmia-related genetic variants with phenotypes documented in electronic medical records. *JAMA* 2016;315:47–57.

20. Biesecker LG. Long QT syndrome and potentially pathogenic genetic variants. *JAMA* 2016;315:2467–2468.
21. Amendola LM, Jarvik GP, Leo MC, et al. Performance of ACMG-AMP variant-interpretation guidelines among nine laboratories in the clinical sequencing exploratory research consortium. *Am J Hum Genet* 2016;99:247.
22. Garber KB, Vincent LM, Alexander JJ, et al. Reassessment of genomic sequence variation to harmonize interpretation for personalized medicine. *Am J Hum Genet* 2016;99:1140–1149.
23. Harrison S, Dolinsky J, Knight-Johnson VL, et al. Clinical laboratories collaborate to resolve variant interpretations submitted to ClinVar. *Genet Med*; e-pub ahead of print 16 March 2017.
24. ClinVar. ClinVar Database, October 2016 version. Available at: ftp://ftp.ncbi.nlm.nih.gov/pub/clinvar/xml/archive/2016//ClinVarFullRelease_2016-10.xml.gz. Accessed 31 October 2016.
25. Lek M, Karczewski KJ, Minikel E, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536:285–291.
26. Karczewski KJ, Weisburd B, Thomas B, et al. The ExAC browser: displaying reference data information from over 60,000 exomes. *Nucleic Acids Res* 2017;45:D840–D845.
27. Exome Aggregation Consortium. ExAC Database version 0.3. Available at: ftp://ftp.broadinstitute.org/distribution/ExAC_release/release0.3/ExAC.r0.3.sites.vcf.gz. Accessed 17 January 2017.
28. Kobayashi Y, Yang S, Nykamp K, Garcia J, Lincoln SE, Topper SE. Pathogenic variant burden in the ExAC database: an empirical approach to evaluating population data for clinical variant interpretation. *Genome Med* 2017;9:13.
29. Song W, Gardner SA, Hovhannisyan H, et al. Exploring the landscape of pathogenic genetic variation in the ExAC population database: insights of relevance to variant classification. *Genet Med* 2016;18:850–854.
30. Walsh R, Thomson KL, Ware JS, et al. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. *Genet Med* 2017;19:192–203.
31. ClinVar Representation of Clinical Significance in ClinVar and Other Variation Resources at NCBI, 2017. <https://www.ncbi.nlm.nih.gov/clinvar/docs/clinsig/>. Accessed 8 March 2017.
32. Amberger JS, Bocchini CA, Schiettecatte F, et al. OMIM.org: online mendelian inheritance in man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res* 2015;43:D789–D798.
33. Spurdle AB, Healey S, Devereau A, et al. ENIGMA—evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. *Hum Mutat* 2012;33:2–7.
34. Thompson BA, Spurdle AB, Plazzer JP, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet* 2014;46:107–115.
35. Elmore JG, Longton GM, Carney PA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* 2015;313:1122–1132.
36. Sprague BL, Conant EF, Onega T, et al. Variation in mammographic breast density assessments among radiologists in clinical practice: a multicenter observational study. *Ann Intern Med* 2016;165:457–464.
37. Myriad Genetics Analyst and Investor Day Presentation, September 2015. http://files.shareholder.com/downloads/MYGN/3639646904x0x661835/b6aec217-2edd-4afc-be8e-d0a36414ce91/AnalystDay_Presentation_FINAL.pdf. Accessed 12 January 2017.
38. Cook-Deegan R, Conley JM, Evans JP, et al. The next controversy in genetic testing: clinical data as trade secrets? *Eur J Hum Genet* 2013;21:585–588.
39. Angrist M, Cook-Deegan R. Distributing the future: the weak justifications for keeping human genomic databases secret and the challenges and opportunities in reverse engineering them. *Appl Transl Genom* 2014;3:124–127.
40. MyriadPro Terms of Use, 2017. <https://www.myriadpro.com/company/terms-of-use/>. Accessed 17 January 2017.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

© The Author(s) 2017