Clinically impactful differences in variant interpretation between clinicians and testing laboratories: a single-center experience

Austin Bland, BSc[1], Elizabeth A. Harrington, MSc[1], Kyla Dunn, MSc[1,2], Mitchel Pariani, MSc[1,2], Julia C.K. Platt, MSc[1,2], Megan E. Grove, MSc[1,2] and Colleen Caleshu, MSc[1,2]

Purpose: To describe the frequency and nature of differences in variant classifications between clinicians and genetic testing laboratories.

Methods: Retrospective review of variants identified through genetic testing ordered in routine clinical care by clinicians in the Stanford Center for Inherited Cardiovascular Disease. We compared classifications made by clinicians, the testing laboratory, and other laboratories in ClinVar.

Results: Of 688 laboratory classifications, 124 (18%) differed from the clinicians' classifications. Most differences in classification would probably affect clinical care of the patient and/or family (83%, 103/124). The frequency of discordant classifications differed depending on the testing laboratory ($P < 0.0001$) and the testing laboratory's classification ($P < 0.00001$). For the majority (82/124, 66%) of discordant classifications, clinicians were more conservative (less likely to classify a variant pathogenic or likely pathogenic). The clinicians' classification was discordant with one or more submitter in ClinVar in 49.1% (28/57) of cases, while the testing laboratory's classification was discordant with a ClinVar submitter in 82.5% of cases (47/57, $P = 0.0002$).

Conclusion: The clinical team disagreed with the laboratory's classification at a rate similar to that of reported disagreements between laboratories. Most of this discordance was clinically significant, with clinicians tending to be more conservative than laboratories in their classifications.

Genet Med advance online publication 14 December 2017

Key Words: cardiovascular genetics; genetic counseling; genetic testing; interpretation; variant classification

INTRODUCTION

The number of available genetic tests is rapidly increasing, as is the number of genes included in any given test.[1] At the same time that clinicians are handling a much larger volume of variants, it has become clear that classifying variants is more challenging than previously thought.[2] Insights from large-scale sequencing efforts such as gnomAD have called into question historical variant classification standards and have led to the reclassification of many variants.[3,4] Data-sharing efforts such as ClinVar have revealed that disagreement about the classification of variants is not infrequent.[1,5–10] These developments have sparked changes in variant classification approaches, including a shift to greater stringency.[6,7,11]

There is evidence that the frequency of disagreements in classifications is particularly high in cardiovascular genetics.[7–9] The 2015 American College of Medical Genetics and Genomics variant classification guidelines were designed to be broadly applicable to all areas of genetics.[11] The lack of gene and disease specificity has made application of these guidelines in specialty areas such as cardiology problematic.[12] In response to these challenges, many clinicians in cardiovascular genetics have taken a more active role in the classification of variants. The majority (81%) of clinical cardiovascular genetic counselors report that their clinical teams reevaluate variant classifications received from genetic testing laboratories.[13]

While previous studies have compared classifications by different laboratories, differences in classifications made by laboratories and clinicians have not been examined. We aimed to delineate the frequency and nature of differences in variant classifications between cardiovascular genetics clinicians and genetic-testing laboratories.

MATERIALS AND METHODS

The Stanford Center for Inherited Cardiovascular Disease provides interdisciplinary cardiovascular genetics care for cardiomyopathies, arrhythmias, dyslipidemias, and aortopathies. The data set included all variants reviewed by clinicians in our center following clinical genetic testing between May 2007 and August 2016. Study methods are outlined in **Supplementary Figure S1** online.

Clinicians' classifications

As part of routine clinical care, members of the clinical team reviewed and classified variants reported via clinical genetic

testing by accredited genetic-testing laboratories. The clinical team's variant classification process is detailed in **Supplementary Figure S2**.

Comparison of the clinicians' and laboratory's classifications

Classifications of variants by clinicians were compared with those reported by the genetic-testing laboratory on the most recently issued clinical genetic-test report. To facilitate such comparisons, all classifications were translated into the American College of Medical Genetics and Genomics (ACMG) five-tier system (pathogenic (P), likely pathogenic (LP), variant of uncertain significance (VUS), likely benign (LB), benign (B)).[11]

Clinician–laboratory classification pairs were categorized as discordant if the clinicians' classification and the laboratory's classification differed in terms of a major classification category (for example, P/LP versus VUS) (**Supplementary Figure S3a**). Classification pairs were considered concordant if classifications were exactly the same (for example, LB versus LB) or differed only in the degree of confidence (for example, LP versus P).

We also sought to determine whether classification differences would affect clinical care of the patient or family. Clinically significant discordance was defined as P/LP versus VUS/LB/B, since P/LP variants are likely to be used to direct care while VUS/LB/B typically are not (**Supplementary Figure S3b**). In defining clinically significant discordance we took into consideration the fact that many clinicians treat P and LP variants in the same way.

ClinVar comparisons

For the subset of variants with discordant classifications, the clinicians' and the testing laboratory's classifications were compared to classifications submitted to ClinVar by other clinical genetic-testing laboratories. Data in ClinVar were accessed on 5 November 2016.

RESULTS

Summary of data set

The data set included a total of 639 unique variants with 688 laboratory variant classifications (**Table 1**). Fourteen genetic-testing laboratories were represented (**Supplementary Table S1**), though 95% of classifications came from the five most frequently used laboratories.

Discordance between clinicians and testing laboratory

The clinicians' and testing laboratory's variant classifications were discordant 18% (126/688) of the time (**Figure 1a**). Discordance that would affect clinical care was found in 83% (105/126) of discordant classifications and 15% (105/688) of all classifications (**Figure 1a**). In most cases discordance arose because the clinical team was more conservative; clinicians were less likely than the laboratory to consider a variant pathogenic or likely pathogenic ($P < 0.00001$; **Figure 1b**). The most frequently observed difference was discordance

Table 1 Variant characteristics

Total variants observed	639	
Total variant classifications	Laboratory	Clinicians
	688	639
Pathogenic	222 (32%)	71 (11%)
Likely pathogenic	68 (10%)	124 (19%)
Variant of unknown significance	379 (55%)	431 (67%)
Likely benign	10 (2%)	1 (<1%)
Benign	9 (1%)	12 (2%)
Genes	98	
MYBPC3	104 (15%)	
MYH7	98 (14%)	
KCNQ1	38 (6%)	
RYR2	27 (4%)	
SCN5A	25 (4%)	
TTN	23 (3%)	
Testing laboratories	14	
Novel variants	307 (45%)	
Missense variants	551 (81%)	
Segregation data available	202 (29%)	

Testing laboratories are listed in **Supplementary Table S1 online**. Some variants were reported in results from multiple laboratories, resulting in more laboratory classifications than variants.

between a laboratory classification of LP/P and a clinician classification of VUS, which occurred in 78% (82/105) of clinically significant instances of discordance (**Figure 1b**).

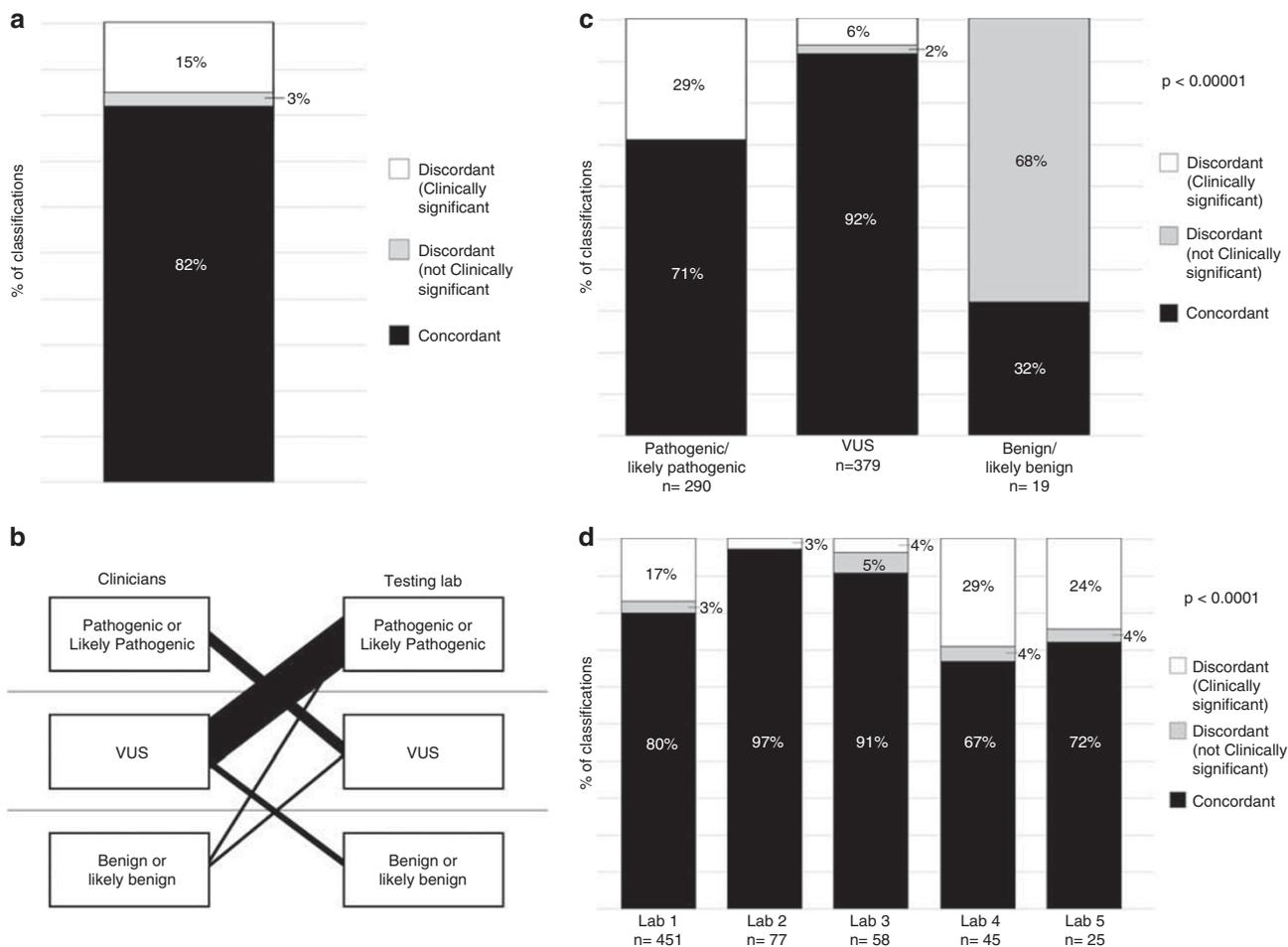## Discordance varies by laboratory classification and laboratory

The frequency of discordance varied depending on the testing laboratory's classification (**Figure 1c**). Discordance was less common when the testing laboratory defined the variant as a VUS (8%) compared with L/LP (29%) and B/LB (68%) ($P < 0.0001$). Discordance also varied depending on the testing laboratory used, ranging from 3% to 33% in the five most commonly used genetic-testing laboratories ($P < 0.00001$) (**Figure 1d**).

Of note, variant characteristics such as type of variation, novelty, case count, gene, or segregation data were not associated with discordance. We saw a bivariate association between timing of testing and frequency of discordance. However, this was confounded by changes in laboratories used over the course of the study period. On multivariate logistic regression, including both timing of testing and testing laboratory, timing variables were not significant.

ClinVar comparisons and discrepancies

Fifty-seven of the variants with discordant classifications in our data set had a classification by at least one other clinical genetic testing laboratory in ClinVar.

We first sought to assess the level of discordance for these variants among genetic-testing laboratories in ClinVar (including the testing laboratory). Over half of the discordant

Figure 1 Discordance between classifications by cardiovascular genetics clinicians and the genetic testing laboratory. (**a**) Frequency of discordance between clinicians and testing laboratory. (**b**) Nature of discordance. Classification pairs noted as clinicians:lab—VUS:LP/P (82/125, 65.6%); LP/P:VUS (23/125, 18.4%); VUS:LB/B (13/125, 10.4%); LB/B:VUS (6/125, 4.8%); LB/B:LP/P (1/125, 0.8%), LP/P:LB/B (0/125, 0%); *P* < 0.0001 for Fisher's exact test across clinicians–laboratory classification pairs. (**c**) Frequency of discordance varies by testing laboratory's classification. *P* < 0.0001 for Fisher's exact test across three testing-lab classification categories. (**d**) Frequency of discordance varies by testing laboratory. VUS = variant of uncertain significance. *P* < 0.0001 for Fisher's exact test across the top five most frequently used laboratories.

variants with entries in ClinVar from more than one genetic-testing laboratory (25/45, 55.6%) had discordant classifications between laboratories in ClinVar, and most of these were clinically significant (88.0%, 22/25).

We then compared the classifications by the clinical team and by the testing laboratory to the classifications by other clinical genetic testing laboratories provided in ClinVar. The clinical team's classification was discordant with one or more laboratory in ClinVar in 49.1% (28/57) of cases. The testing laboratory's classification was discordant with another laboratory in ClinVar in 82.5% of cases (47/57, *P* = 0.0002). More laboratories in ClinVar agreed with the clinical team's classification than with the testing laboratory's classification (mean 1.4 vs. mean 0.6, *P* = 0.0005).

DISCUSSION

Clinicians' variant classifications differed from the classification reported by the genetic-testing laboratory 18% of the

time, and most of these differences (83%) would affect clinical care for the patient and/or family (15% of all classifications). Clinicians tended to be more cautious in their classifications, calling fewer variants LP/P. Notably, the clinical team's classifications were more in line with classifications by other laboratories, as represented in ClinVar.

It is impossible to know which classification is correct, the clinicians' or the laboratory's. Variant classification is inherently probabilistic and to date most methods remain subjective. Without a gold standard or a method for knowing the ground truth regarding a variant's pathogenicity, we cannot assess which classification is accurate. It is notable that when the clinicians' classification differed from the laboratory's classification, the clinicians' classification was more in line with classifications by other laboratories. It is also important to consider the time frame for the study (2007–2016) when interpreting our finding that the most common form of disagreement was when the laboratory classified

a variant as LP/P and the clinicians classified it as a VUS. During this time period variant classifications across the field of genetics have become more stringent.[14] It is likely that if these variants were reassessed now, some of them would be reclassified by the laboratory as VUS, thus resolving the discordance in classification.

The frequency of disagreement between classifications by the clinical team and laboratory classifications that we observed (18%) is comparable to the frequency of disagreement observed between genetic-testing laboratories (12–53%) in previous studies.[1,5–9] Discordance between laboratories arises from a variety of factors, most notably from lack of data sharing, outdated classifications, and inconsistencies in the use of classification criteria.[5–7,10] Recent efforts have shown that when these key issues are addressed and classifications are discussed on an "as needed" basis, discordance can often be resolved.[5–7,10] However, it is notable that even with these efforts, 13–44% of discordant classifications cannot currently be resolved.[5–7,10] These data point to the probabilistic nature of variant classification and they also suggest that a small portion of variants is particularly challenging or controversial. We found that variants with discordance in our data set had a particularly high level of discordance between clinical genetic-testing laboratories in ClinVar (55.6% compared to 17–19% across all of ClinVar).[1,8] This suggests that at least some of the variants that were discordant in our data set are especially challenging to classify.

The majority of cardiovascular genetics groups assess variant classifications after receiving a genetic test report, motivated by a sense of responsibility for the result as well as previous experiences with reclassifications and disagreements between laboratories.[13] These motivations are in line with studies that suggest discordance in genetic-test interpretation is more common in cardiovascular genetics.[7–9] They also conform with our finding that clinicians approach variant classification with caution; when they disagree with the testing laboratory they tend to downgrade a LP/P classification to a VUS. A certain amount of caution in response to challenges in variant classification may be particularly apt in the setting of cardiovascular genetics, where misclassifying a variant as LP/P could lead to inappropriately discharging variant-negative healthy relatives from cardiac screening and leaving them at risk of sudden death. This tendency of clinicians to be less likely to classify a variant as LP/P is also consistent with recent increases in stringency in variant classification practices in the field as a whole.[14]

In addition, our data show that when the clinicians' classification differed from the testing laboratory's classification, more laboratories in ClinVar agreed with the clinicians' classification than with the testing laboratory's classification. This indicates that the clinicians' classification is aligned with the rest of the field. Our data suggest that there is some benefit from clinician involvement in variant classification, and that the expertise of specialized clinicians may be a more important factor in variant interpretation than has been previously established. Periodic reassessment of classifications provides a key opportunity for clinicians to participate in variant classification, whether by performing reassessments themselves or by prompting the testing laboratory to do so. Automatic periodic reassessment of variants is not currently among the services provided by genetic-testing laboratories; however, data from several studies demonstrate that clinically significant reclassifications occur when variants are reevaluated.[6,7,15] When considering the role of clinicians in variant assessment, it is important to note that not all clinicians have the necessary training or expertise to confidently and successfully classify variants. Genetic testing may be ordered by cardiologists who do not have the needed expertise in variant classification to go beyond the assessment provided by the laboratory. Guidelines recommend that genetics evaluations for inherited cardiovascular disease be performed by centers with appropriate expertise and staffing, including cardiovascular genetic counselors as part of the multidisciplinary team.[16,17] Finally, the not insignificant amount of time required for clinicians to assess variant classifications must be considered, particularly with the increasing demand for genetics services and the need to improve efficiency and increase clinical volumes.

While discordance in genetic test interpretation has only recently been broadly recognized, discordance in clinical interpretation within the medical field as a whole is not novel.[8] Examples include differences in interpretation of ventricular tachycardia on electrocardiograms and assessment of anatomic pathology slides.[18,19] Genetics clinicians and laboratories can learn from insights and processes developed in other fields, such as the use of mandatory second-opinion review in surgical pathology.[19] Such review could be performed by the clinical team if they have sufficient expertise, by another genetic testing laboratory, or, more passively, by checking ClinVar. As we learn more about how to minimize, manage, and resolve differences in variant classification, our field will undoubtedly uncover insights and develop standards that could benefit other areas of medicine, such as the movement toward data sharing.

While the uncertainties of variant classification are challenging for clinicians to manage, the ultimate impact is, of course, on patients and families. Of utmost importance are the ways in which clinicians help families navigate the uncertainties inherent in genetic-test interpretation, inconsistencies in classification, and changes in classification over time.[20]

CONCLUSION

Within our cardiovascular genetics center we found that our clinical team's classifications of variants differed from the genetic testing laboratory's classifications 18% of the time, with most of those differences impacting clinical care. This rate of disagreement is comparable to the rate of disagreement reported between genetic testing laboratories, reflecting the need for efforts to decrease discordance and increase accuracy in genetic-test interpretation. Our data suggest there may be

a role for expert clinicians in the assessment of variant classification.

SUPPLEMENTARY MATERIAL

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

DISCLOSURE

C.C has been a consultant and/or adviser for Recombine, Invitae, Phosphorus, and MyoKardia, and owns shares in Personalis. K.D. has been a consultant for Invitae. The other authors declare no conflict of interest.

REFERENCES

1. Rehm HL, Berg JS, Brooks LD, et al. ClinGen—the clinical genome resource. *N Engl J Med*. 2015;372:2235–2242.
2. Caleshu C, Ashley EA. Taming the genome: towards better genetic test interpretation. *Genome Med*. 2016;(8)1:9–11.
3. Lek M, Karczewski KJ, Eric V, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nat Publ Gr*. 2016;536:285–291.
4. Andreasen C, Nielsen JB, Refsgaard L, et al. New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. *Eur J Hum Genet*. 2013;21:918–928.
5. 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;98:1067–1076.
6. Garber KB, Vincent LM, Alexander JJ, Bean LJH, Bale S, Hegde M. Reassessment of genomic sequence variation to harmonize interpretation for personalized medicine. *Am J Hum Genet*. 2016;99:1140–1149.
7. Harrison SM, Dolinsky JS, Knight Johnson AE, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet Med*. 2017. 14;19:1096–1104.
8. 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;0(March):1–9.
9. Balmana 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.
10. Furqan A, Arscott P, Girolami F, et al. Care in specialized centers and data sharing increase agreement in HCM genetic test interpretation. *Circ Cardiovasc Genet*. 2017;10:e001700.
11. 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;17:405–423.
12. Kelly MA, Caleshu C, Morales A, 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*. (in press).
13. Reuter C, Grove ME, Orland K, Spoonamore K, Caleshu C. Abstract: Clinical cardiovascular genetic counselors take a leading role in team-based variant interpretation. *J Genet Couns*. 2016;25:1347–1472.
14. Rehm HL. A new era in the interpretation of human genomic variation. *Genet Med*. 2017;19:1092–1095.
15. Das K J, Ingles J, Bagnall RD, Semsarian C. Determining pathogenicity of genetic variants in hypertrophic cardiomyopathy: importance of periodic reassessment. *Genet Med*. 2014;16:286–293.
16. Ackerman MJ, Priori SG, Willems S, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies: this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Heart Rhythm*. 2011;8:1308–1339.
17. Hershberger RE, Lindenfeld J, Mestroni L, Seidman CE, Taylor MRG, Towbin JA. Genetic evaluation of cardiomyopathy—a Heart Failure Society of America practice guideline. *J Card Fail*. 2009;15:83–97.
18. 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*. 2017;165:457–464.
19. Manion E, Cohen MB, Weydert J. Mandatory second opinion in surgical pathology referral material: clinical consequences of major disagreements. *Am J Surg Pathol*. 2008;32:732–737.
20. Ingles J, Semsarian C. Conveying a probabilistic genetic test result to families with an inherited heart disease. *Heart Rhythm*. 2014;11:1073–1078.