Care in Specialized Centers and Data Sharing Increase Agreement in Hypertrophic Cardiomyopathy Genetic Test Interpretation

Title: Care in Specialized Centers and Data Sharing Increase Agreement in HCM Genetic Test Interpretation

Manuscript number: CIRCCVG/2016/001700

Author(s): Colleen Caleshu, Stanford University
Aisha Furqan, California State University
Patricia Arscott, University of Michigan Health System
Francesca Girolami, Careggi University Hospital
Allison L. Cirino, Brigham and Women's Hospital, Boston, MA
Michelle Michels, Erasmus Medical Center
Sharlene Day, University of Michigan
Iacopo Olivotto, Referral Center for Cardiomyopathies, Careggi University Hospital
Carolyn Ho, Brigham and Women's Hospital

Euan Ashley, Stanford University School of Medicine

Eric Green, MyoKardia

Care in Specialized Centers and Data Sharing Increase Agreement in HCM Genetic Test Interpretation

Running title: Furqan: Reducing genetic test interpretation discordance

Aisha Furqan[1], Patricia Arscott[2], Francesca Girolami[3], Allison L Cirino[4], Michelle Michels[5], Sharlene M Day[2], Iacopo Olivotto[6], Carolyn Y Ho[4], Euan Ashley[7,8], Eric M. Green[9], SHaRe Consortium, Colleen Caleshu[7,10]

1. California State University, Stanislaus, CA, USA
2. University of Michigan, Division of Cardiovascular Medicine, Ann Arbor, MI, USA
3. Genetics Unit, Careggi University Hospital Florence, Italy
4. Cardiovascular Division, Brigham and Women's Hospital, Boston, MA, USA
5. Department of Cardiology, Thoraxcenter, Erasmus Medical Center, Rotterdam, the Netherlands.
6. Referral Center for Myocardial Diseases, Azienda Ospedaliera Universitaria Careggi, Florence, Italy
7. Stanford Center for Inherited Cardiovascular Disease, Stanford, CA, USA
8. Division of Cardiovascular Medicine, Stanford Medical Center, Stanford, CA, USA
9. MyoKardia Inc., South San Francisco, CA, USA
10. Division of Medical Genetics, Stanford Medical Center, Stanford, CA, USA

Address for Correspondence:
Colleen Caleshu
Stanford Center for Inherited Cardiovascular Disease
Falk CVRC, 300 Pasteur Drive, Stanford, CA 94305-5406
Phone: 650-725-6273
Fax: 650-498-7452
Email: ccaleshu@stanfordhealthcare.org

Word count: 6345 (total); Journal Subject Terms: Genetics, Cardiomyopathy, Diagnostic Testing, Hypertrophy

ABSTRACT

Background: Clinically impactful differences in the interpretation of genetic test results occur among laboratories and clinicians. To improve the classification of variants, a better understanding of why discrepancies occur and how they can be reduced is needed.

Methods and results: We examined the frequency, causes, and resolution of discordant variant classifications in the Sarcomeric Human Cardiomyopathy Registry (SHaRe), a consortium of international centers with expertise in the clinical management and genetic architecture of hypertrophic cardiomyopathy (HCM). Of the 112 variants present in patients at >1 center, 23 had discordant classifications between centers (20.5%, Fleiss' kappa 0.54). Discordance was more than twice as frequent among clinical laboratories in ClinVar, a public archive of variant classifications (315/695 variants, 45.2%, Fleiss' kappa 0.30; p<0.001). Discordance in SHaRe most frequently occurred because HCM centers had access to different privately held data when making their classifications (75.0%). Centers reassessed their classifications based on a comprehensive and current data summary, leading to reclassifications that reduced the discordance rate from 20.5% to 10.7%. Different interpretations of allele frequencies and co-occurrence with pathogenic variants contributed to residual discordance.

Conclusions: Discordance in variant classification between HCM centers is largely attributable to privately held data. Some discrepancies are due to differences in expert assessment of conflicted data. Discordance was markedly lower among centers specialized in HCM than among clinical laboratories, suggesting that optimal genetic test interpretation occurs in the context of clinical care delivered by specialized centers with both clinical and genetics expertise.

KEY WORDS: Genetic testing, hypertrophic cardiomyopathy, genetic counseling, genetic test interpretation

INTRODUCTION

Hypertrophic cardiomyopathy (HCM) is an inherited cardiovascular disease characterized by left ventricular hypertrophy that occurs in the absence of pressure overload, systemic disease, or infiltrative processes. Individuals with HCM are at increased risk for adverse clinical events including heart failure, atrial fibrillation, stroke, and sudden cardiac death[1]. Disease-causing sarcomere variants are identified in a third of HCM cases with another 15% having an inconclusive genetic test result[2]. Genetic testing for HCM has become routine in centers specialized in the disease and is recommended in multiple medical guidelines[3,4]. Once a variant is identified on genetic testing, a variety of data points are reviewed and an assessment is made as to the likelihood that the variant causes HCM[5]. This leads to a classification that the variant likely causes disease (pathogenic, likely pathogenic), is inconclusive (variant of uncertain significance), or is unlikely to cause disease (likely benign, or benign). The primary benefits of genetic testing arise when a pathogenic or likely pathogenic variant is found, which can help in establishing a definitive diagnosis in the patient and in assessing risk of disease in healthy relatives.

At the same time that genetic testing for cardiovascular diseases like HCM has become common practice, the complexities of interpreting such tests and the need for reliable and consistent standards for interpretation have become increasingly evident. Large-scale population sequencing datasets such as ExAC (Exome Aggregation Consortium, exac.broadinstitute.org) have demonstrated that rare variation is abundant in the genome, challenging the assumption that most rare variation causes severe Mendelian genetic disease and questioning the pathogenicity of thousands of specific variants[6]. Data sharing efforts like ClinVar have also revealed challenges in current variant classification approaches[7]. ClinVar, a public repository of variant classifications submitted by clinical laboratories and researchers, has facilitated comparisons between laboratories, revealing

that differences in interpretation are not uncommon. Many of these differences are clinically impactful; one laboratory may classify a variant as pathogenic prompting the clinician to use that variant in diagnostic evaluations and to assess risk in healthy relatives, while another laboratory calls it a variant of uncertain significance and as such it would not be used in clinical care. The frequency of differences in interpretation between laboratories has ranged from 17-53% in different studies[7–10]. These shifts in the field have revealed the need for improved approaches to genetic test interpretation.

Efforts are underway to both resolve disagreements between laboratories and to improve genetic test interpretation guidelines to increase both agreement and accuracy [5,7,11]. Within cardiology specifically, the Cardiovascular Domain Working Group of the ClinGen initiative is developing gene and disease specific variant interpretation guidance[7] (https://www.clinicalgenome.org/working-groups/clinical-domain/sub-groups/cardiovascular/).

A better understanding of why disgreements in classification occur and how they can be resolved will aid efforts to improve variant classification strategies and guide clinicians in navigating the clinical implications of differences in interpretation. To gain such insights we investigated the frequency, origins, and resolution of disagreements in variant classifications among centers specialized in HCM participating in the Sarcomeric Human Cardiomyopathy Registry (SHaRe, http://www.thesharegistry.org).

METHODS

SHaRe is an international consortium that amalgamates de-identified patient-level data on inherited cardiomyopathies from established institutional datasets at participating centers. At the time of analysis, SHaRe contained clinical and genetic testing data on 4944 patients with HCM, from the following centers: Stanford University (STU), Brigham and Women's Hospital (BWH), University of Michigan (UMH), Erasmus University (ERA), and

Careggi University (FLO). All centers have expertise in both the clinical management of HCM and comprehensive genetics evaluations, including family evaluations and interpretation of genetic testing. SHaRe centers assigned a classification to each sarcomere variant present in their population based on both the interpretation provided by the genetic testing laboratory that performed testing, as well as the center's judgment.

Variant data in eight sarcomere genes (*ACTC1, MYBPC3, MYH7, MYL2, TNNI3, TNNT2,* and *TPM1)*, were downloaded from the SHaRe database (March 2015). To set the SHaRe data in context, we also examined discordant classifications in ClinVar, using data on the same genes (downloaded April 2015). To focus on clinical laboratories, ClinVar submissions from OMIM and research laboratories were excluded. ClinVar submitters included in this analysis: Laboratory for Molecular Medicine, GeneDx, LabCorp, Blueprint Genetics, Children's Hospital of Eastern Ontario, Invitae, University of Washington, Emory Genetics Laboratory, Genetic Services Laboratory, University of Chicago, Neurogenetics Laboratory (Royal Perth Hospital).

Any individual variant that was seen by more than one HCM center or clinical laboratory had the potential to be classified discordantly. A variant was considered to have discordant classifications if the classifications from two or more groups crossed a major classification category (i.e. likely pathogenic/pathogenic vs. variant of uncertain significance; likely pathogenic/pathogenic vs. likely benign/benign, variant of uncertain significance vs. likely benign/benign). Classifications that differed only by degree of confidence within the same major classification category were considered concordant (i.e. likely pathogenic vs. pathogenic, likely benign vs. benign). The frequency of discordant classifications was calculated by dividing the number of variants with discordant classifications by the total number of variants with classifications by more than one group (Figure 1). Fleiss' kappa was used to assess inter-rater reliability, a modification of Cohen's kappa for three or more reviewers[12].

We determined if disagreements in classification were clinically significant, meaning they would impact medical care, such as diagnosis in the patient or use of predictive genetic testing for healthy at-risk family members. Discordance was considered clinically significant if it involved a likely pathogenic or pathogenic classification and any other classification (i.e. likely pathogenic/pathogenic vs. variant of uncertain significance or likely pathogenic/pathogenic vs. likely benign/benign).

The reasons for discordance in SHaRe were assessed by comparing the rationale provided by each SHaRe center to justify their classification of that variant. These data was available for 20/23 discordant variants (the other three became concordant upon the centers' review of their initial classification).

To assist in resolution of discordance among SHaRe centers, each center was provided with up-to-date summaries of all available data on each discordant variant reported by their center and asked to reassess their classification and provide a rationale. To assess why discordance remained after these reclassifications we compared centers' rationales for their final classification and examined the data available on each variant. This included an assessment of the number of data points suggesting the variant may be benign, which included co-occurrence with another likely pathogenic or pathogenic variant in >1% of cases[2], presence in reference samples with MAF > 0.00004[13], failure to segregate, and occurrence with other phenotypes.

RESULTS

Discordance in SHaRe is lower than in ClinVar

Participants in SHaRe had 589 unique sarcomere variants, of which 112 were seen by more than one center (Figure 1). Discordant classifications were present in 23 of these variants (20.5%; Fleiss' kappa 0.54, 95% CI 0.38-0.69). To contextualize this rate of discordance we compared it to the rate of discordance among clinical laboratories in

ClinVar. ClinVar contains 2,405 sarcomere variants, of which 695 variants were submitted by more than one laboratory and 314 were discordant (45.2%; Fleiss' kappa 0.36, 95% CI 0.30-0.42; p <0.001 for the comparison between SHaRe and ClinVar discordance). In both ClinVar and SHaRe, most discordant classifications were clinically significant (SHaRe: 19/23, 82.6%; ClinVar: 229/314, 72.9%; p=0.75).

Discordance is often due to lack of data sharing and outdated data

Comparison of the rationale for initial variant classifications provided by SHaRe centers revealed that most variants had more than one reason for discordance (mean 2.5, standard deviation 1.4). The most common reason for discordance was differential access to privately held data (15/20, 75%), from either the SHaRe center's clinical experience (12/20, 60%) or the genetic testing laboratory's internal data (12/20, 60%) (Figure 2). This most frequently involved co-occurrence of the discordant variant with another pathogenic variant, suggesting the discordant variant may be benign. This occurred for 11/20 discordant variants (55%); in four of those cases that data was held only by a SHaRe center, in another four it was held only by the laboratory that did the testing, and in the remaining three it was held by both the testing laboratory and SHaRe center. For 7/20 discordant variants (35%) the SHaRe centers differed in their access to segregation data. In five cases the segregation data was privately held by the SHaRe center, in two cases it was held by the laboratory that did the genetic testing and in no cases was it held by both the testing laboratory and the SHaRe center. Data that was not available to all SHaRe centers included several types of data that suggest the variant could be benign: being seen with inconsistent phenotypes (4/20, 20%), cases being from the same ancestry without ancestry matched controls (1/20, 5%), and presence in reference samples (1/20, 5%). Consistent with the impact of differential access to data, more than half of the HCM cases associated with the discordant variants were not publicly available; 125 cases were

published or publicly available while 193 were only available in a private dataset that not all centers initially had access to (Table 1).

In some cases the centers disagreed in their classifications because the publicly available data that they used differed, including citing different literature (9/20, 45%), predictions from in silico models (5/20, 25%), population frequencies (3/20, 15%), and assessments of evolutionary conservation (3/20, 15%) (Figure 2). This sometimes occurred because the SHaRe centers' initial classifications were done at different times, so some centers had classified the variant using data that was now outdated.

For nearly a third of discordant variants SHaRe centers cited one or more data point that was identical, but was interpreted differently by different centers (Figure 2). For example, for p.Gly490Arg (c.1468G>A) in *MYBPC3,* three sites were aware the variant had been seen in cases that had another variant that was deemed pathogenic. One site used that to reach a variant of uncertain significance classification while the other two sites classified the variant as likely pathogenic or pathogenic despite that data.

Partial resolution of discordance can be achieved through data sharing

When SHaRe centers were asked to provide their rationale for their initial classification to help illuminate sources of discordance, three centers changed initial classifications based on review of the data the center already had on the variant in light of their current approach to classifications. This resolved discordance for three of the 23 (13%) initially discordant variants (Figure 3A, Supplemental Table 1).

To resolve the remaining discordance in SHaRe, we compiled up-to-date comprehensive summaries of the data on the 20 remaining discordant variants, including both publically available data and data privately available to each center (Table 1). Each SHaRe center was asked to review a detailed narrative summary of this data and provide an updated classification for their discordant variants. This reduced discordance further, from 23/112 initially (20.5%) to 12/112 (10.7%) (Figure 3). Nearly all of the remaining

discordant classifications were clinically significant (11/12, 91.7%). Most of these were variant of uncertain significance vs. likely pathogenic (7/12, 58.3%) or pathogenic (2/12, 8.3%) (Table 1, Supplemental Table 1). Two were likely benign or benign vs. likely pathogenic (2/12, 16.7%). For seven of the twelve variants that remained discordant, at least one center changed their classification, yet that reclassification did not resolve discordance (Supplemental Table 1). There were no reclassifications in the other five variants.

To gain insight into why discordance was not completely resolved despite the SHaRe centers having access to the same data, we examined the data gathered on the 20 discordant variants the centers were asked to re-assess (Table 1). We also compared HCM centers' rationales for their final classifications. Note that complete data on rationale for final classifications was only available for 11 of 12 variants that remained discordant.

Among the variants that reached concordance, none of the variants reclassified to likely pathogenic or pathogenic had evidence suggesting they may be benign. In contrast, all of the variants reclassified to likely benign or benign had evidence suggesting they may be benign, with a mean of 2.7 types of benign evidence per variant. The variants that remained discrepant had a mean of 1.8 types of benign evidence per variant, suggesting the data on these variants was more conflicted and did not point as clearly towards a benign or pathogenic classification.

Notably, in nearly two-thirds of variants that remained discordant at least one center remarked in their rationale that they suspect the variant is a modifier (7/11, 63.6%). Consistent with this, these variants had features typically associated with modifying variants; most of these variants had co-occurred with a pathogenic variant (9/12, 75%) and were present in reference samples (11/12, 91.7%). This contrasts to the variants with resolved discordance in which a minority had co-occurred with a pathogenic variant (3/8, 37.5%) and only half had been seen in reference samples (4/8, 50.0%). Among the

variants that were seen in references samples, the mean minor allele frequency was higher for those with resolved discordance (0.11, three of four reclassified to (likely) benign) than those with unresolved discordance (0.0021). It is also notable that 8 of 11 variants with resolved discordance were missense while all variants with unresolved discordance (12/12) were missense, consistent with greater challenges in classifying missense variation and their potential role as modifiers.

Examining the rationales that centers provided for their final classifications, two areas of disagreement occurred in over half of variants that remained discordant: differing assessments of whether the variant was sufficiently rare in reference samples (7/11 variants, 63.6%) and differing interpretations of how co-occurrence with another pathogenic variant affected classification (7/11 variants, 63.6%).

DISCUSSION

As genetic testing for inherited cardiovascular conditions such as HCM has become common place, there has been increasing awareness of the complexity of genetic test interpretation and the not infrequent occurrence of clinically impactful differences in the classification of variants. While prior studies have examined differences in interpretation among laboratories, the current study dissects differences in the interpretations used by clinical centers, where genetic testing data is translated into patient care. Our data show disagreements in classification are far less frequent within the setting of specialized HCM centers with expertise in disease management, phenotypic and family assessment, and understanding of the genetic architecture of HCM.

The initial rate of disagreement in variant classification among SHaRe centers (20.5%) is at the lower range of the rate of disagreement among laboratories reported to date (17-53%)[7–10] and is less than half that seen in ClinVar for the same set of sarcomere genes (45.2%) (Figure 3C). Furthermore, half of disagreements among HCM centers were

resolved via sharing of comprehensive up to date data. The rate of discordance in SHaRe after efforts to resolve disagreements, 10.7%, is the lowest yet reported. These data suggest that the complexities of genetic test interpretation are best addressed within the context of specialized centers leveraging the benefits of data-sharing. Consistent with our finding that discordance is lower when genetic testing occurs in the context of a specialized center, professional societies have recommended that genetics evaluations for heritable cardiomyopathies be carried out in such specialized centers[4].

The lower rate of discordance among HCM centers as compared to clinical laboratories could have several different origins, including selection of patients for genetic testing, benefits of comprehensive family-based genetic evaluations, and application of expertise in HCM. The patients included in SHaRe all have clear diagnoses of HCM. This is in contrast to the sample of patients who undergo genetic testing at clinical laboratories, which includes those with clear diagnoses as well as those with borderline or questionable diagnoses. Prior studies in various genetic conditions have shown that the yield of genetic testing is lower in patients referred to clinical laboratories than in studies of patients with firm diagnoses[14,15]. The current study demonstrates that not only is yield lower in the heterogeneous samples seen in the clinical laboratory setting, but discordance is higher. The comprehensive genetics evaluations provided by HCM centers is another potential sources of lower discordance among those centers. The genetic evaluations performed by these centers goes beyond genetic testing on the index patient to include analysis of 3-4 generation pedigrees, phenotyping and genotyping of family members, and expert assessment for genocopies and phenocopies. The impact of these evaluations is evident in the effect that data generated by individual centers had on classifications and discordance. Centers' classifications arose not only from the data provided by the genetic testing laboratory or published in the literature but from the centers' own clinical evaluations, such as segregation analyses performed by the center. The lower rate of

discordance among HCM centers may also arise from application of these centers' expertise in disease-specific genetic variation. In their 2015 guidelines on sequence variant interpretation, the American College of Medical Genetics and Genomics specifically pointed to the need for gene-disease specific classification criteria[5]. Prior studies have demonstrated the impact of gene-disease expertise on variant classification. Comparing variant classifications by a laboratory specialized in connective tissue disorders to other laboratories, Pepin et al (2015) found that in a third of cases the other laboratory failed to factor in key aspects of protein structure and function that significantly impact classification[16]. Amendola et al (2016) found that differences in gene-disease expertise contributed to differences in classification among laboratories in the Clinical Sequencing Exploratory Research (CSER) consortium[9]. Thus the gene-disease expertise of specialized HCM centers along with the comprehensive genetics evaluations they perform, and the selection of patients for genetic testing may all contribute to optimization of genetic test interpretation and minimization of discordance.

Our data also speak to the importance of data sharing, both between clinicians and laboratories regarding an individual case and between individuals laboratories or centers and the broader community. The most frequent reason for discordance within SHaRe was lack of data sharing at the point of initial classification; the majority of classification differences occurred, at least in part, because SHaRe centers had access to differing private datasets. Data-sharing is particularly critical for a disease like HCM that is characterized by such marked genetic heterogeneity; 56% of variants found on HCM genetic testing by one laboratory were seen in just a single family[2]. This makes it challenging for any one laboratory or center to accumulate enough data to determine the appropriate classification for such variants, particularly when they are missense. Data-sharing efforts like SHaRe, ClinVar, and gnomAD (Genome Aggregation Database, http://gnomad.broadinstitute.org/, formerly ExAC) are shifting variant classifications and

impacting clinical care[6]. There is debate about whether it should be mandatory for the ordering clinician to share clinical data with the laboratory and the laboratory to share data with the community through efforts such as ClinVar. Our data suggests that such sharing would improve variant classifications and have clinical benefit.

A subset of discordance in SHaRe (13%) was resolved simply by the center reconsidering the data they already had on the variant in light of current knowledge and classification practices. An additional 35% of discordance was resolved by review of current up to date data. This underscores the importance of ongoing re-review of variant interpretations. Similarly, Das et al (2014) found that re-review of the pathogenic, likely pathogenic, and uncertain variants in their HCM center lead to clinically impactful reclassifications in 10% of their patients with variants[17]. The vast majority of laboratories do not consistently re-review variants, typically only doing so if they observe the variant again. Given the marked genetic heterogeneity in HCM this means that it may be many years before a variant is re-reviewed and a substantial subset will never be re-reviewed. Yet both our data and prior studies[17,18] show that re-review can lead to changes in classification that impact medical care in an appreciable subset of variants.

Despite basing revised variant classifications on identical data, 10.7% of variants seen by more than one center in SHaRe remained discordant. This residual discordance appears to be attributable to differences in expert opinion when the available data are subjective and conflicted, particularly whether the variant is sufficiently rare and whether it is seen too often in tandem with a pathogenic variant. Further pointing to the impact of differences in expert opinion is our observation that a third of initial discordance in SHaRe was at least partially attributable to differing interpretations of the same data. It is possible that some of the residual discordance could be resolved by agreeing upon and using identical classification criteria, such as cut offs for rarity and co-occurrence. Work is underway within the ClinGen Cardiovascular Domain Working Group to develop disease

and gene specific guidance on matters like rarity and co-occurence[7] (https://www.clinicalgenome.org/working-groups/clinical-domain/sub-groups/cardiovascular/). These guidelines will be informed by insights into rarity of pathogenic variation for a given gene and disease provided by analyses of large disease and reference datasets[13] and frequency of co-occurrence in large disease cohorts[2]. While such guidelines will undoubtedly improve variant interpretation, they may not completely resolve discordance; Amendola et al (2016) found that even when laboratories used identical classification criteria and identical data, discordance remained in a substantial subset of variants[9]. Another possible explanation for the unresolved discordance in SHaRe is suggested by the fact that in nearly two-thirds of variants that remained discordant at least one center suspected the variant was a modifier. This is a class of variation that is poorly understood and is not accounted for in existing classification guidelines. Given the limitations of our current knowledge and guidelines, experts sometimes need to make judgment calls in interpreting variant data and as such a certain amount of discordance will remain due to differences in expert opinion. Moreover, discordant interpretation of clinical data is not unique to genetic testing. Comparable rates of disagreement have been reported across a range of medical specialties and tests including assessment of ventricular tachycardia on ECGs (22%)[19], subtyping of sarcoma on histopathology (27%)[20], and assessment of wall motion abnormalities on dobutamine stress tests (15%)[21]. Such data has led some authors to recommend routine second opinions for some medical tests[22]. Data sharing via efforts like ClinVar is now allowing for a passive form of second opinion, in which laboratories and clinicians can check ClinVar to see how other groups classify a variant.

Clinical implications:

Discordance in the classification of variants from DNA sequencing data occurs both between genetic testing laboratories and between clinical centers. Most of the disagreements in classification in both SHaRe and ClinVar would affect clinical management of the patient and/or family, highlighting the importance of the current challenges with variant interpretation. This raises the question of how clinical teams should be handling variant classifications, particularly when there are disagreements in classification. Concerns regarding these differences and a sense of responsibility for the clinical impact of the test interpretation have led many clinical cardiovascular genetics groups to start making their own assessments of variants received on clinical genetic testing[23]. In many ways this parallels how these teams often handle cardiac testing; in addition to reviewing reports for cardiac imaging they also look closely at the primary imaging data and interpret them independently. The need for periodic re-review of classifications suggested by our data and prior studies may provide further justification for clinicians taking greater responsibility for ongoing genetic test interpretation, since such re-review is currently not common practice among laboratories[17,18]. Finally, the greater level of agreement among HCM centers than among clinical laboratories suggests genetic testing for HCM is best done in the setting of an expert center.

CONCLUSION

Disagreement in the interpretation of genetic test results exists among genetic testing laboratories and clinical HCM centers. Moreover, most of these disagreements are of sufficient magnitude to impact clinical utilization of the test results. The majority of disagreements are due to privately held or outdated data, which can be ameliorated by increased data sharing and periodic re-review of currently available data. However, differences in expert assessment of complex data will continue to be a source of discordance for the near future. Notably, the discordance between centers with expertise

in HCM management and genetic testing was significantly lower than that seen between clinical genetic testing laboratories. These findings highlight the important benefits that can be achieved when expertise in disease management and family evaluations is combined with expertise in genetic interpretation.

ACKNOWLEDGEMENTS

We thank James Signorovitch for statistical consultation.

FUNDNG:

The SHaRe consortium is funded by an unrestricted research grant from MyoKardia

CONFLICT OF INTEREST DISCLOSURE:

Colleen Caleshu: Consultant, Advisor – Recombine; Advisor – Invitae; Advisor – Phosphorous.

Eric Green: Employee of and owns shares in MyoKardia, Inc.

Euan Ashley: Ownership interest, Advisor – Personalis Inc, Advisor - SequenceBio

Aisha Furqan, Patricia Arscott, Francesca Girolami, Allison L Cirino, Michelle Michels, Sharlene M Day, Iacopo Olivotto, Carolyn Y Ho: None

REFERENCES

1. Maron BJ, Casey S a, Poliac LC, Gohman TE, Almquist a K, Aeppli DM. Clinical course of hypertrophic cardiomyopathy in a regional United States cohort. *JAMA*. 1999;281:650–655.
2. Alfares AA, Kelly MA, McDermott G, Funke BH, Lebo MS, Baxter SB, et al. Results of clinical genetic testing of 2,912 probands with hypertrophic cardiomyopathy: expanded panels offer limited additional sensitivity. *Genet Med*. 2015;17:880–888.
3. Gersh B, BJ Maron, RO Bonow, JA Dearani, MA Fifer, MS Link, et al. 2011 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy. *J Thorac Cardiovasc Surg*. 2011;142:e153–e203.
4. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, 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;8:1308–39.

5. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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.
6. Lek M, Karczewski K, Minikel E, Samocha K, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285–91.
7. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, et al. ClinGen - The Clinical Genome Resource. *N Engl J Med*. 2015;372:2235–42.
8. Balmana J, Digiovanni L, Gaddam P, Walsh MF, Joseph V, Stadler ZK, 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.
9. Amendola LM, Jarvik GP, Leo MC, McLaughlin HM, Akkari Y, Amaral MD, 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.
10. Amendola LM, Dorschner MO, Robertson PD, Salama JS, Hart R, Shirts BH, et al. Actionable exomic incidental findings in 6503 participants: Challenges of variant classification. *Genome Res*. 2015;25:305–315.
11. 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.

12. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol. Bull. 1971;76:378–382.
13. Whiffin N, Minikel E, Walsh R, O'Donnell-Luria A, Karczewski K, Ing AY, et al. Using high-resolution variant frequencies to empower clinical genome interpretation. *bioRxiv*. 2016. 10.1101/073114.
14. Napolitano C, Priori SG, Schwartz PJ, Bloise R, Ronchetti E, Nastoli J, et al. Genetic testing in the long QT syndrome: development and validation of an efficient approach to genotyping in clinical practice. *JAMA*. 2005;294:2975–2980.
15. Kapplinger JD, Tester DJ, Salisbury BA, Carr JL, Harris-Kerr C, Pollevick GD, et al. Spectrum and prevalence of mutations from the first 2,500 consecutive unrelated patients referred for the FAMILION long QT syndrome genetic test. *Hear Rhythm*. 2009;6:1297–1303.
16. Pepin MG, Murray ML, Bailey S, Leistritz-Kessler D, Schwarze U, Byers PH. The challenge of comprehensive and consistent sequence variant interpretation between clinical laboratories. *Genet Med*. 2015;1–5.
17. 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.
18. Aronson SJ, Clark EH, Varugheese M, Baxter S, Babb LJ, Rehm HL. Communicating new knowledge on previously reported genetic variants. *Genet Med*. 2012;14:713–719.
19. Herbert ME, Votey SR, Morgan MT, Cameron P, Dziukas L. Failure to agree on the electrocardiographic diagnosis of ventricular tachycardia. *Ann Emerg Med*. 1996;27:35–38.
20. Presant BCA, Russell WO, Alexander RW, Fu YS. Soft-Tissue and Bone Sarcoma Histopathology Peer Review: The Frequency of Disagreement in Diagnosis and the

Need for Second Pathology Opinions The Southeastern Cancer Study Group Experience. *J Clin Oncol*. 1986;4:1658–1661.

21. Paetsch I, Jahnke C, Ferrari VA, Rademakers FE, Pellikka PA, Hundley WG, et al. Determination of interobserver variability for identifying inducible left ventricular wall motion abnormalities during dobutamine stress magnetic resonance imaging. *Eur Heart J*. 2006;27:1459–1464.

22. Al-Ibraheemi A, Folpe AL. Voluntary Second Opinions in Pediatric Bone and Soft Tissue Pathology: A Retrospective Review of 1601 Cases From a Single Mesenchymal Tumor Consultation Service. *Int J Surg Pathol*. 2016;24:685–691.

23. 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.

TABLES AND FIGURES:

Table 1. Discordant variants in SHaRe

A. Variants with Discordance Resolved

Gene	Variant	Reclass-ification	Benign evidence count*	Unrelated HCM cases	Other phenotypes	Seen with LP/P variant	Segregation		Presence in reference samples	
							Meioses segregating	Meioses failing to segregate	Highest MAF	Population with highest MAF
MYBPC3	**p.Gln998Glu (c.2992C>G)**	VUS to LB/B	2	35			1	0	0.09	Latino (ExAC)
MYBPC3	**p.Ser217Gly (c.649A>G)**	VUS to LB	4	7	DCM, SIDS	3/7	0	2	0.012	South Asian (ExAC)
MYBPC3	**p.Val189Ile (c.565G>A)**	VUS to LB	2	8		3/8	0	0	0.0043	South Asian (ExAC)
MYBPC3	**p.Pro371Arg (c.1112C>G)**	LP to VUS	1	4		4/4	3	0	0	
MYBPC3	**c.927-9G>A**	VUS to P	0	30			5	0	0	
MYBPC3	**c.1224-2A>G**	VUS to LP	0	1			0	0	0	

MYH7	**p.Arg1420Trp (c.4258C>T)**	VUS to LP	0	11		0	0	0.000015	European (ExAC)

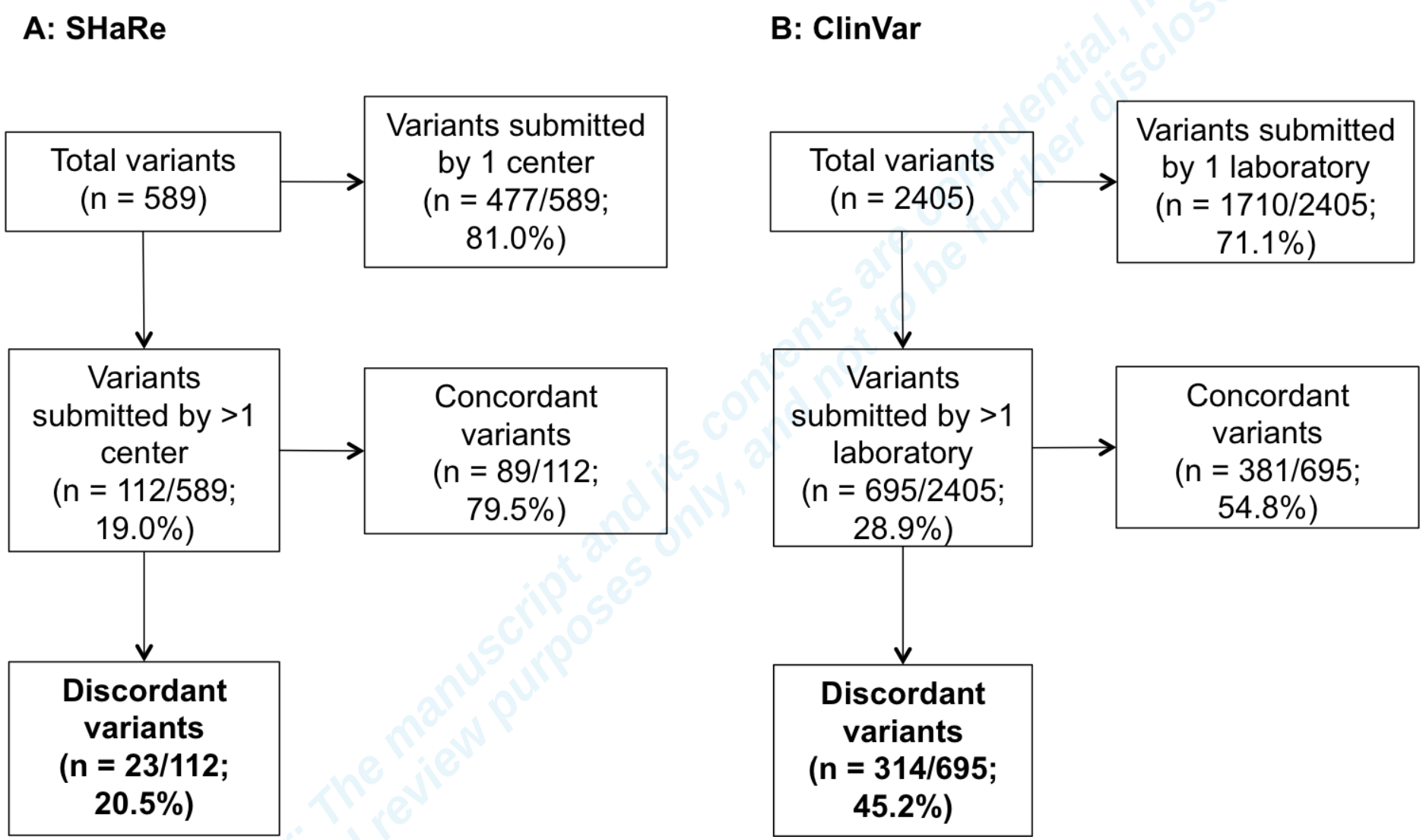
B. Variants with Discordance Unresolved

Gene	Variant	Remaining discordance	Benign evidence count*	Unrelated HCM cases	Other phenotypes	Segregation: Seen with LP/P variant	Segregation: Meioses segregating	Segregation: Meioses failing to segregate	Presence in reference samples: Highest MAF	Presence in reference samples: Population with highest MAF
MYBPC3	**p.Glu619Lys (c.1855G>A)**	LP v LB	3	10	DCM, LVNC, WPW	4/10	1	0	0.0013	European (ExAC)
MYH7	**p.Met982Thr (c.2945T>C)**	LP v B	3	19	DCM, increased LVWT, SCD with dilatation	7/19	1	0	0.0013	European (ExAC)
MYH7	**p.Asn1327Lys (c.3981C>A)**	VUS v LB	2	14		2/12	1	0	0.018	Ashkenazi (LMM)
MYH7	**p.Lys1459Asn (c.4377G>T)**	LP v VUS	3	15	Ebstein's, Brugada	2/9	1	0	0.00051	European (ExAC)

MYH7	**p.Arg1606Cys (c.4816C>T)**	LP v VUS	0	2			0	0	0.0000077	European (ExAC)
MYH7	**p.Arg204His (c.611G>A)**	LP v VUS	1	20		2/20	1	0	0	
MYBPC3	**p.Arg810His (c.2429G>A)**	LP v VUS	2	29		5/27	5	0	0.00006	European (ExAC)
MYBPC3	**p.Arg1002Gln (c.3005G>A)**	LP v VUS	2	3	DCM, giant RA & arrhythmia		1	0	0.00012	East Asian (ExAC)
MYBPC3	**p.Gly531Arg (c.1591G>C)**	LP v VUS	1	3		1/3	2	0	0.000031	European (ExAC)
MYBPC3	**p.Gly490Arg (c.1468G>A)**	LP v VUS	2	10		4/8	2	0	0.00045	Finnish (ExAC)
TNNT2	**p.Arg278Cys (c.832C>T)**	P v VUS	2	51	DCM	5/47	8	0	0.0016	Other (ExAC)
MYH7	**p.Thr1377Met (c.4130C>T)**	P v VUS	0	25			0	0	0.000029	European (pooled)

Summary data on SHaRe discordant variants that became concordant after reassessment (A) and those that remained discordant (B). VUS = variant of uncertain significance, LP = likely pathogenic, P = pathogenic, LB = likely benign, B = benign. MAF = minor allele frequency. ExAC = Exome Aggregation Consortium (exac.broadinstitute.org). DCM = dilated cardiomyopathy, SIDS = sudden infant death syndrome, LVNC = left ventricular non-compaction, WPW = Wolff-Parkinson-

White, LVWT = left ventricular wall thickness, SCD = sudden cardiac death, RA = right atrium. * Note that comprehensive data was not gathered on the three variants that became concordant when the HCM centers reviewed the rationales for their initial classifications.



508

509 **Figure 1. Assessment of discordance in SHaRe and ClinVar**. Variants for 8 sarcomere genes were downloaded from SHaRe
510 (A) and ClinVar (B). Variants with classifications from >1 SHaRe center or ClinVar submitter were identified. Classifications were
511 compared across centers or submitters to assess discordance.
512

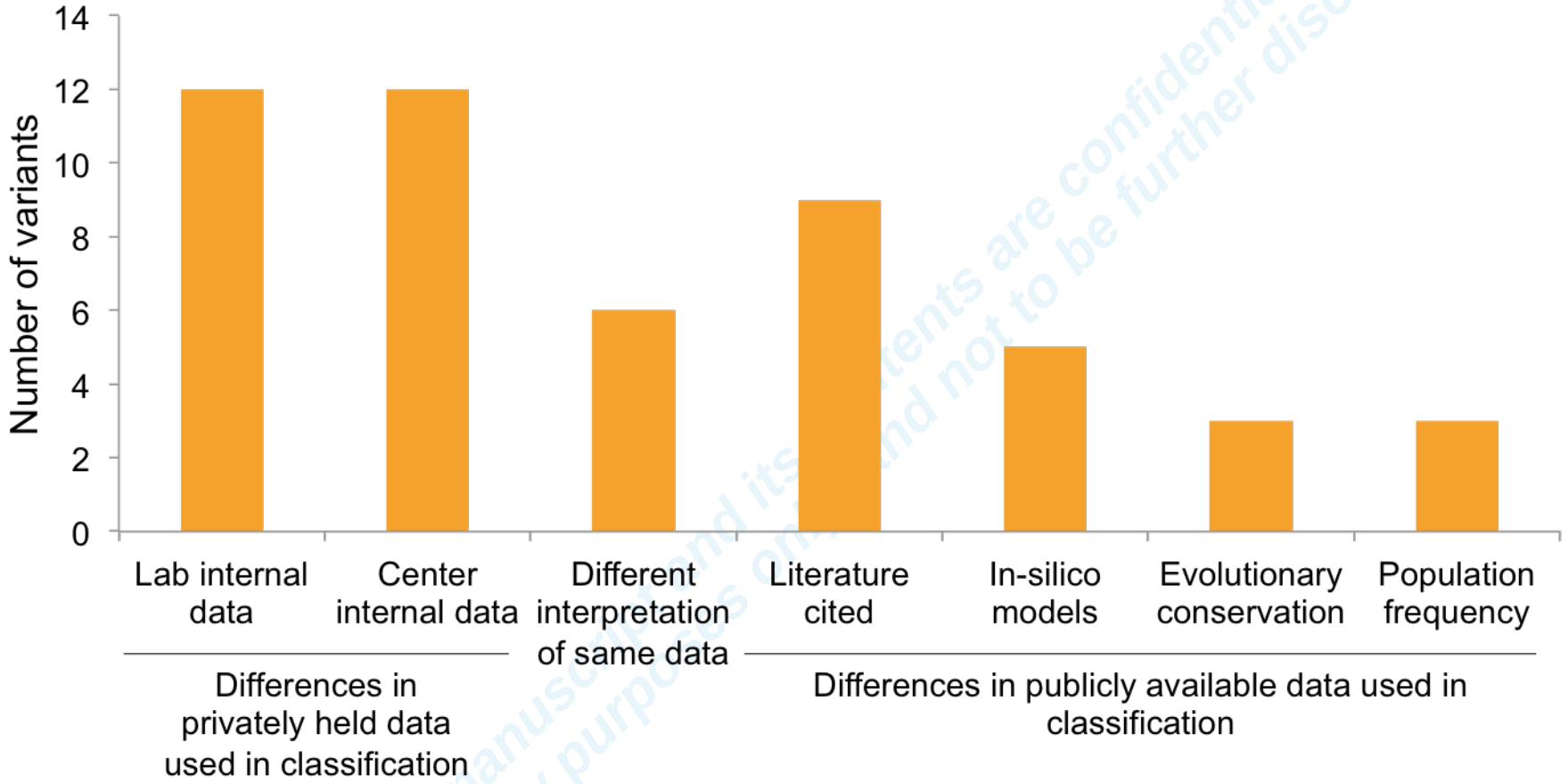
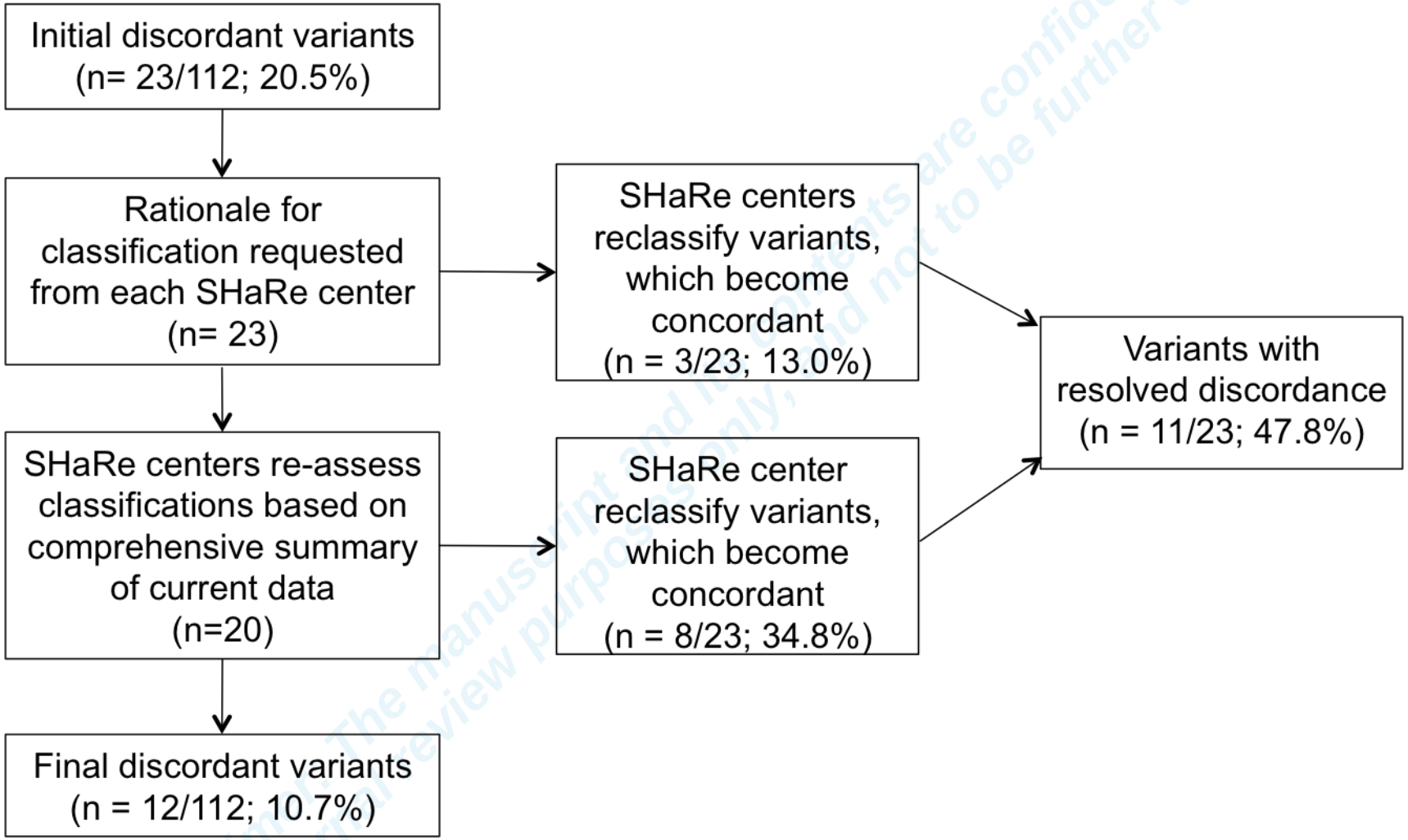
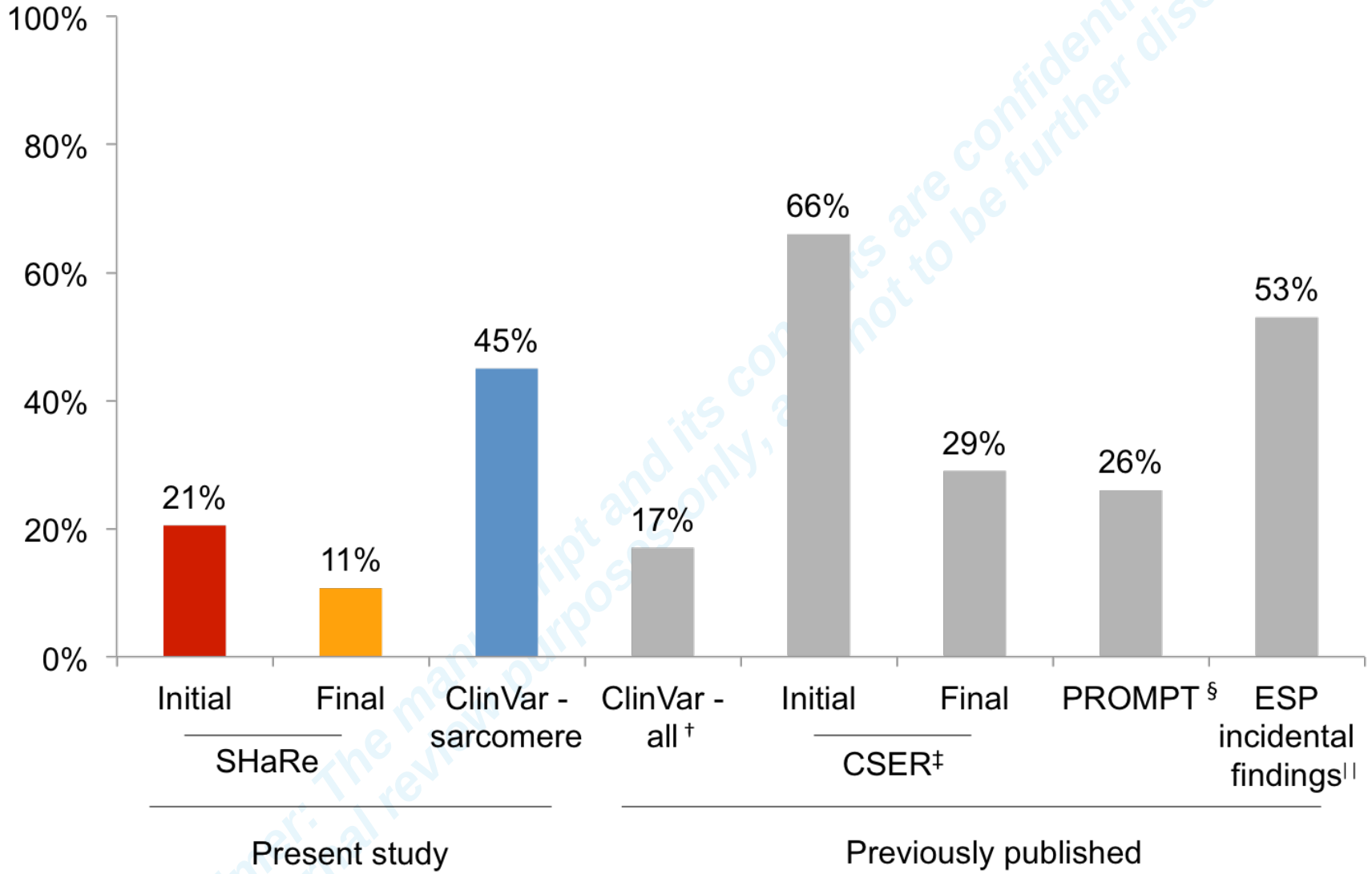


Figure 2. Reasons for initial discordance among SHaRe centers. Reasons for discordance were assessed by comparing the rationales for each center's classifications.

A: Resolution of discordance in SHaRe



C: Discordance Rates in SHaRe, ClinVar, and Prior Publications



519

Figure 3. Reclassification of SHaRe variants and resolution of discordance. A: Process of reclassification of variants and (partial) resolution of discordance in SHaRe. *When SHaRe centers pulled their rationale for their initial classification, three centers changed their classification given their current classification methods. **B:** Initial and final classifications in SHaRe, shown for variants that became concordant after reclassification (left) and those that remained discordant (right). Each line represents one center's classifications of one variant. Centers are designated by line color (see legend). **C:** Discordance rates in SHaRe and ClinVar from the present study and previously published discordance rates. †Discordance in ClinVar across all genes[7]. ‡Discordance in classification of select variants studied by CSER (Clinical Sequencing Exploratory Research Consortium), before and after efforts to reduce discordance[9]. §Discordance rate among clinical laboratories on variants in cancer genes submitted to PROMPT (Prospective Registry of Multiplex Testing)[8]. ||Discordance between reviewers of potentially actionable incidental findings in ESP (Exome Sequencing Project)[10]. VUS = variant of uncertain significance, LP = likely pathogenic, P = pathogenic, LB = likely benign, B = benign. SHaRe centers: Stanford University (STU), Brigham and Women's Hospital (BWH), University of Michigan (UMH), Erasmus University (ERA), and Careggi University (FLO).

Supplemental Table 1. initial and final classifications of discordant variants in SHaRe

A. Discordance resolved

Gene	Variant	BWH		FLO		ERA		STU		UMH	
		Initial	Final	Initial	Final	Initial	Final	Initial	Final	Initial	Final
MYBPC3	**p.Val189Ile (c.565G>A)**	LB	LB					VUS	LB	LB	LB
MYBPC3	**p.Ser217Gly (c.649A>G)**	LB	LB					VUS	LB	LB	B
MYBPC3	**c.927-9G>A**	P	P			VUS	P	LP	P	P	P
MYBPC3	**p.Pro371Arg (c.1112C>G)**	VUS	VUS	LP	VUS						
MYBPC3	**c.1224-2A>G**					VUS	LP	LP	LP		
MYBPC3	**p.Asp605Asn**	VUS	VUS							P	VUS

	(c.1813G>A)*								
MYBPC3	**p.Ser858Asn**	LP	VUS			VUS	VUS		
	(c.2573G>A)*								
MYBPC3	**p.Gln998Glu**					VUS	LB	B	B
	(c.2992C>G)								
MYH7	**p.Arg1420Trp**			P	LP			VUS	LP
	(c.4258C>T)								
TNNT2	**p.Trp287***	LP	LP	VUS	P	P	P		
	(c.860G>A)*								
TPM1	**p.Glu192Lys**	LP	LP			VUS	LP		
	(c.574G>A)								

536

537 **B. Discordance unresolved**

Gene	Variant	BWH		FLO		ERA		STU		UMH	
		Initial	Final	Initial	Final	Initial	Final	Initial	Final	Initial	Final
MYBPC3	**p.Gly490Arg**			LP	LP	P	LP			VUS	VUS

	(c.1468G>A)								
MYBPC3	**p.Gly531Arg**	VUS	VUS	LP	LP				
	(c.1591G>C)								
MYBPC3	**p.Glu619Lys**	VUS	LB	LP	LP			LB	LB
	(c.1855G>A)								
MYBPC3	**p.Arg810His**	VUS	VUS	LP	LP	VUS	VUS	VUS	LP
	(c.2429G>A)								
MYBPC3	**p.Arg1002Gln**	VUS	VUS	P	LP				
	(c.3005G>A)								
MYH7	**p.Arg204His**	VUS	VUS	VUS	P	LP	LP	LP	LP
	(c.611G>A)								
MYH7	**p.Met982Thr**	VUS	LB	LP	LP			VUS	B
	(c.2945T>C)								
MYH7	**p.Asn1327Lys**	LB	LB			VUS	VUS	VUS	VUS
	(c.3981C>A)								
MYH7	**p.Thr1377Met**	VUS	VUS	P	P			LP	LP

	(c.4130C>T)										
MYH7	**p.Lys1459Asn**			VUS	LP	P	LP	LP	VUS	VUS	VUS
	(c.4377G>T)										
MYH7	**p.Arg1606Cys**					VUS	LP			LP	VUS
	(c.4816C>T)										
TNNT2	**p.Arg278Cys**	VUS	VUS	LP	P	P	LP	LP	VUS	P	VUS
	(c.832C>T)										

538 The 23 variants in SHaRe with discordant classifications. Each center's initial and final classification is noted. VUS = variant of
539 uncertain significance, LP = likely pathogenic, P = pathogenic, LB = likely benign, B = benign. SHaRe centers: Stanford University
540 (STU), Brigham and Women's Hospital (BWH), University of Michigan (UMH), Erasmus University (ERA), and Careggi University
541 (FLO). *Discordance in these three variants was resolved when the centers reviewed their initial classifications and reclassified
542 these variants based on the data they already possessed (Figure 3A). The remaining variants were re-assessed based on review of
543 all data currently available to the authors.