Molecular Diagnosis & Therapy
https://doi.org/10.1007/s40291-025-00784-8

ORIGINAL RESEARCH ARTICLE

Making Sense of Missense: Benchmarking MutScore for Variant Interpretation in Inherited Cardiac Diseases

Alessandra Pia Porretta[1,2] · Véronique Fressart[3] · Elodie Surget[1] · Charles Morgat[1] · Adrien Bloch[3] · Anne Messali[1] · Vincent Algalarrondo[1] · Géraldine Vedrenne[1] · Etienne Pruvot[2] · Antoine Leenhardt[1] · Isabelle Denjoy[1] · Fabrice Extramiana[1]

Accepted: 27 April 2025

Abstract

Background Accurate interpretation of genetic variants still represents a major challenge. According to current recommendations from the American College of Medical Genetics and Genomics (ACMG), variant interpretation relies on a comprehensive analysis including, among others, computational data for prediction of variant pathogenicity. However, the predictive accuracy of in silico tools is often limited, and results are frequently inconsistent. In the current study, we evaluated the predictive performance of a previously described innovative classifier (MutScore) for missense variants in our cohort of probands with inherited cardiac diseases (InCDs).

Methods We retrospectively reviewed missense variants detected in our cohort of probands with InCDs. Variants were analyzed with four in silico tools commonly used in our diagnostic pipelines (CADD, Polyphen-2, Alpha-missense and Revel) and with MutScore, a new meta-predictor combining data on variant location with the output of 16 existing predictors. For each variant, we recorded the original classification (established according to scientific evidence available at the time of molecular diagnosis) and the updated classification performed at the present time, according to ACMG standards.

Results We detected 252 missense variants in our cohort of 517 patients affected by InCDs. MutScore was the most proficient tool in classifying variants (0.89 maximum area under the curve [95% confidence interval (CI) 0.85–0.94]). Compared to Revel, the second-best predictor, MutScore showed superior sensitivity (73% vs 57%) at the maximum tolerated false-positive rate of 10%, higher specificity (0.83 vs 0.36) and a markedly lower false-positive rate (0.17 vs 0.64), supporting a more nuanced and accurate assessment, especially for benign or likely benign variants. MutScore also appeared to perform better for variants located in genes associated with channelopathies than for variants in cardiomyopathy-related genes. Notably, when comparing the original and updated classification, 27% (69/252) of missense variants underwent a change in classification over the 9-year follow-up period. Among these, reclassification had a significant impact on clinical management in one third of cases (i.e., variants of uncertain significance upgraded to pathogenic or likely pathogenic variants or vice versa), with a 4.8% increase in molecular diagnosis of InCDs over the 9-year period.

Conclusion Our study supports the excellent performance of MutScore in a real-life dataset of missense variants associated with the rare subset of InCDs. MutScore represents a promising application of artificial intelligence with major potential in cardiogenetics to improve diagnostic precision in clinical practice. In addition, our results highlight the importance of periodic reanalysis of variants, incorporating newly available scientific evidence, as attested by the significant implications for patient management and clinical decision-making.

✉ Alessandra Pia Porretta
Alessandra-pia.porretta@chuv.ch

[1] CNMR Maladies Cardiaques Héréditaires Rares, APHP, Hôpital Bichat Claude-Bernard, Paris, France

[2] Service of Cardiology, Heart and Vessel Department, Centre Hospitalier Universitaire Vaudois (CHUV), Rue du Bugnon 46, 1011 Lausanne, Switzerland

[3] Service de Biochimie Métabolique, Groupe Hospitalier Pitié-Salpêtrière, Paris, France

1 Introduction

Over recent decades, the development of next-generation sequencing (NGS) technologies has contributed to making genetic testing more timely and economically viable, facilitating the broader accessibility of genomic sequencing. However, a considerable gap persists between sequencing of genomes and accurate interpretation of genetic variants [1, 2]. Indeed, even for classic monogenic diseases,

Published online: 21 May 2025

Key Points

MutScore is an innovative meta-predictor that represents an application of artificial intelligence to enhance the in silico prediction of variant effects.

In our dataset of missense variants detected in a cohort of probands with inherited cardiac diseases, we demonstrated that MutScore outperformed four commonly used in silico predictors (CADD, Polyphen-2, Alpha-missense and Revel) in classifying variants.

MutScore has significant potential in cardiogenetics to improve the diagnostic management of patients and their families.

By demonstrating that 27% of missense variants underwent reclassification over the 9-year follow-up period, our results also highlight the importance of periodic variant reanalysis in clinical practice, incorporating newly available scientific evidence.

the definition of variant clinical significance is often hampered by different levels of complexity, including genetic/allelic heterogeneity (i.e., the"one disease–multiple genes"paradigm), incomplete penetrance and variable expressivity as well as the possible involvement of the same gene in different diseases (i.e., the “one gene–many diseases” paradigm) [1, 3, 4]. In this context, the detection of variants of uncertain significance (VUS) has become an increasingly frequent finding, which unfortunately lacks clinical actionability [1]. For this reason, accurate initial classification of variants is crucial, and periodic report updates are recommended to promote (based on newly available scientific evidence) variant reclassification, with potential major implications for patient management [1].

According to current standards for the interpretation of sequence variants [1], in silico computational data for predicting variant pathogenicity represent a fundamental necessity in the analysis workflow. They provide a powerful tool to bypass the impracticable number of experimental validations and the prohibitive sample sizes that would be necessary to prove rare variant effects [5]. Previous algorithms predicted pathogenicity based on unidimensional features such as the evolutionary conservation or the biochemical properties of the amino acid substitution. Conversely, newer meta-predictors integrate scores and outputs from multiple prediction tools, using a combinatory approach based on machine learning techniques [6, 7].

However, the robustness of meta-predictors strictly depends on the type of training protocol and on the quality of training data. If trained on “noisy” data, characterized by random fluctuations that are superfluous to model prediction, overfitting may occur and explain the decline in predictive accuracy on novel datasets. Conversely, circularity and information redundancy among meta-predictor individual features may cause the reinforcement of redundant rather than independent data, leading to overprediction. Training data representativeness is then a key point to warrant the reliability of calculator predictions. In this context, it remains uncertain whether meta-predictors maintain consistent prediction accuracy when applied on gene/variant types that are underrepresented in training datasets.

In the present study, we first evaluated and compared the predictive performance of a new innovative classifier for missense variants (MutScore) to four other prediction algorithms (Polyphen-2, CADD, Revel and Alpha-missense) commonly applied in clinical practice using our real-life dataset of variants associated with the rare subset of inherited cardiac diseases (InCDs). Secondly, we evaluated whether MutScore and the best predictor among the four other algorithms provided consistent performance between channelopathy- versus familial cardiomyopathy-related variants.

2 Materials and Methods

2.1 Real-Life Dataset

Our real-life dataset specifically mirrors variants detected and classified among probands referred for genetic testing to our clinical diagnostic laboratory, between July 2014 and July 2023, due to confirmed or suspected InCDs. These latter included both channelopathies and cardiomyopathies, the clinical diagnosis of which was established, following detailed cardiac assessment and according to international criteria [8, 9], by a team of qualified cardiologists of the Reference Center for Inherited Arrhythmia Syndromes of the Bichat Claude-Bernard University Hospital (Paris, France). All enrolled patients provided written informed consent for genetic testing that allowed personal medical data to be used for research purposes. The study protocol was approved by the appropriate hospital committee and complied with the Declaration of Helsinki 1975 and with its further amendments.

DNA was extracted from blood samples, and genetic analyses were performed using high-throughput sequencing techniques on pre-specified gene panels based on the standard protocols of the Genetics and Cytogenetics Laboratory of the APHP Pitié-Salpêtrière Hospital. Of note, the choice

of gene panel was specifically driven by the patient's clinical phenotype. Each variant identified was systematically validated on a second sample, using capillary sequencing or an alternative technique according to variant type.

Variants were annotated using complementary DNA (cDNA) and the nomenclature of the protein referenced to the Human Gene Mutation Database (HGMD) Professional Refseq transcript. For every detected variant, the original classification (established according to criteria and scientific evidence available at the time of molecular diagnosis) was recorded. Moreover, an updated classification was manually performed at the present time (July 2024) based on current scientific evidence and on the integrated analysis proposed by the American College of Medical Genetics and Genomics (ACMG) [1]. For analysis purposes, we then extracted all missense single-nucleotide variants from the entire dataset.

2.2 In Silico Predictor Benchmarking

The scores of Polyphen-2 [10], CADD [11], Revel [12] and Alpha-missense [13] were annotated using the MobiDetails software (https://mobidetails.iurc.montp.inserm.fr/MD). MobiDetails [14] is a validated online data aggregator developed by the Laboratory of Molecular Genetics of the Montpellier University Hospital, which is currently used in our laboratory for routine variant interpretation in clinical practice. The scores from these four prediction tools are then integrated into our standard pipelines for clinical variant analysis. Accordingly, our study focused on the four mentioned predictors to assess MutScore performance within a real-life diagnostic setting, using tools already applied in everyday clinical workflows. The MutScore was calculated using the freely available software (https://mutscore-wgt7hvakhqew.a.run.app). As previously described [15], the MutLand interface displays data sourced from ClinVar, gnomAD and Uniprot, along with the MutScore result and with the within-gene distribution of clusters of pathogenic/likely pathogenic (PLP) and benign/likely benign (BLB) variants. The MutScore was calculated using the algorithm as previously described [15]. Based on a machine learning approach, MutScore integrates the results of 16 existing unsupervised and independent features with two new scores (positional and amino acid change scores) considering the within-gene topographic distribution of BLB and PLP variants.

Polyphen-2, Revel, Alpha-missense and MutScore provide a score between 0 and 1, while CADD calculates a continuous "phred-scaled" score ranging from 1 to 99. To compare tool-specific variant categorization, two threshold values were applied for each predictor. A variant was defined as BLB when its tool-specific score was below the score enabling the identification of 95% of the BLB variants (i.e., reflecting a 95% sensitivity for BLB variant detection) of the dataset on which the software had been trained. Conversely, a variant was defined as PLP when its tool-specific score was above the value identifying 95% of the PLP variants (i.e., reflecting a 95% sensitivity for PLP variant detection) of the dataset on which the algorithm had been trained. VUS were defined as variants with a score between the lower and upper thresholds. More in detail, the two-sided thresholds for defining BLB and PLP variants correspond to 0.14 and 0.73 for MutScore, and 0.086 and 0.682 for Revel [15].

Predictor benchmarking was performed through the analysis of tool-specific receiver operating characteristic (ROC) curves and the comparison of the area under the curve (AUC). The input represented by the predictor-specific scores was plotted against the last updated and dichotomized variant classification (PLP vs BLB), which, in the absence of a gold-standard for variant interpretation, was used as its most reliable approximation.

Of note, concerning the analysis on reliability for correctly classifying variants, we performed two types of evaluations. We considered that when applying tool-specific thresholds for BLB and PLP variant identification, prediction algorithms do not allow a binary variant classification (i.e., BLB vs PLP) but identify a third "inconclusive" class including variants with a tool-specific score above the threshold identifying BLB variants and below the PLP one. For this reason, we first evaluated the performance metrics (sensitivity, specificity, false-positive/negative rate, etc.) excluding variants with inconclusive tool-specific results (i.e., with a score above the tool-specific threshold identifying BLB variants and below the PLP one). As a further analysis, the same performance metrics were reassessed considering variants with inconclusive scores as false-positive or false-negative results, in order to preserve a binary classification on the same set of variants. As such, the latter approach was performed according to the intention-to-diagnose analysis proposed by Schuetz et al. [16] and Mannhart et al. [17] for the reallocation of diagnostic tests with non-binary results.

2.3 Analysis of Channelopathy- and Cardiomyopathy-Related Variants in the Real-Life Dataset

To assess whether MutScore and the best predictor among the four other algorithms provided consistent performance between channelopathy- and cardiomyopathy-related variants, we extracted from the whole real-life dataset of

missense single-nucleotide variants those located on genes associated with definite or strong evidence of causality to channelopathies or cardiomyopathies according to current recommendations [8]. The channelopathy genes included *KCNQ1*, *KCNH2*, *SCN5 A*, *CALM1*, *CALM2*, *CALM3*, *TRDN*, *KCNE1*, *KCNE2*, *KCNJ2*, *CACNA1 C*, *RYR2*, *CASQ2* and *TECRL*. The cardiomyopathy genes encompassed *MYBPC3*, *MYH7*, *TNNI3*, *TNNT2*, *TPM1*, *ACTC1*, *MYL2*, *MYL3*, *PLN*, *DES*, *FHL1*, *GLA*, *LAMP2*, *PRKAG2*, *PTPN11*, *RAF1*, *TTR*, *ALPK3*, *BAG3*, *FLNC*, *LMNA*, *RBM20*, *TNNC1*, *TTN*, *PKP2*, *DSP*, *DSG2*, *DSC2*, *JUP* and *TMEM43*.

2.4 Statistical Analysis

Categorical variables were expressed as absolute number (*N*) and percentage (%). The Shapiro-Wilk test was used to assess continuous variable distribution; variables with a normal distribution were expressed in terms of mean values ± standard deviation, while non-normal variables were reported as median values and interquartile ranges (p25; p75). The χ^2 test or the Fisher's exact test was performed to assess differences between categorical data. Depending on distribution type (normal or non-normal), the Student's *t* test or the Mann-Whitney test was used to assess differences between continuous variables. All tests were two-sided, and statistical significance was considered for a p value < 0.05. All statistical analysis, including ROC curves and AUC analysis, were performed using the SPSS software package (IBM SPSS Statistics 20.0). The comparison between predictor-specific ROC curves was performed using the MedCalc Software Ltd. (Comparison of AUC of independent ROC curves) based on the method by Hanley and McNeil [18, 19].

3 Results

3.1 Global Overview on the Real-Life Dataset

The cohort of the real-life dataset included 517 probands referred, between 2014 and 2023, for genetic testing due to diagnosed or suspected InCDs. Among them, 221 (42.7%) had no genetic variant detected, 70 (13.5%) presented with at least one non-missense variant (including splicing, truncating and frameshift variants), while at least one missense variant was identified in 226 (43.7%). According to original classification (established at the time of molecular diagnosis), the latter included seven patients with a likely benign (class 2) variant, 126 with a VUS (class 3), 41 probands with a likely pathogenic (class 4) variant and 52 with a pathogenic (class 5) variant. Of note, among the 296 carriers of either a non-missense or a missense variant, 66 were found to carry a second variant (i.e., double-variant carriers), including six non-missense and 60 missense variants. Altogether, we finally annotated 76 non-missense and 286 missense variants.

With the diagnostic yield according to the original variant classification, 302 patients (58.4%) returned a negative genetic result, either for the complete absence of detected variants in 221 probands or due to the identification of non-diagnostic actionable variants (class 1–3 variants) in 81 cases. Conversely, variant detection has been communicated to the remaining 215 patients (41.6%), carrying one or more class 3–5 variants. Interestingly, among the 181 patients identified with VUS, the result was communicated only to 62 patients (34.3%) (Fig. 1).

Focusing on missense variants, the analysis was limited to 252 out of 286 variants after removal of 34 carriers of already annotated variants to avoid duplicates. Among the 252 missense variants and according to the original classification, we reported six class 2 (2.4%), 157 class 3 (62.3%), 40 class 4 (15.9%) and 49 class 5 (19.4%) variants.

Missense variants were globally distributed over all 46 genes (Table 1). Of note, genes associated with definite/strong evidence for channelopathies [8] represented the topographic localization of 55.2% of detected variants. The detailed analysis of gene distribution based on variant classification found that PLP variants were mainly located on genes *TNNI3* (1/1, 100%), *KCNE3* (1/1, 100%), *KCNQ1* 20/25, 80%) and *SCN5 A* (31/45, 68.9%) (Table 1a; see the electronic supplementary material).

Finally, we analyzed variant interpretations comparing the original classification (established according to criteria and scientific evidence available at the time of diagnosis) with the updated classification performed at the present time. We observed that 69 (27.4%) out of 252 missense variants underwent a significant class change over the 9-year follow-up. We observed that 61 (38.9%) out of 157 original VUS were disambiguated as BLB (49/61, 80.3%) or PLP (12/61, 19.7%). Among the 89 original PLP variants, four (4.5%) were downgraded to VUS and four (4.5%) to BLB. Altogether, the last updated classification included 59 BLB variants, 100 VUS and 93 PLP variants (Fig. 2). Of note, 20 (29%) out of 69 variant reclassifications significantly affecting clinical management were then communicated to patients, including VUS upgraded to PLP and PLP variants downgraded to VUS or to BLB, with a 4.8% increase (12 VUS upgraded to PLP out of 252 original missense variants) in molecular diagnosis of InCDs over a 9-year period.

Interestingly, when analyzing the distribution of significant class changes by genes, we observed that 29 (42%) out of 69 reclassifications involved genes associated with definite/strong evidence with channelopathies, among which *RYR2* (12/29), *SCN5 A* (8/29) and *KCNH2* (5/29) were the most frequently represented. Conversely, ten class changes (14.5%) occurred in cardiomyopathy genes, including *LMNA* (4/10), *DSP* (3/10), *DSC2* (2/10) and *DSG2* (1/10).

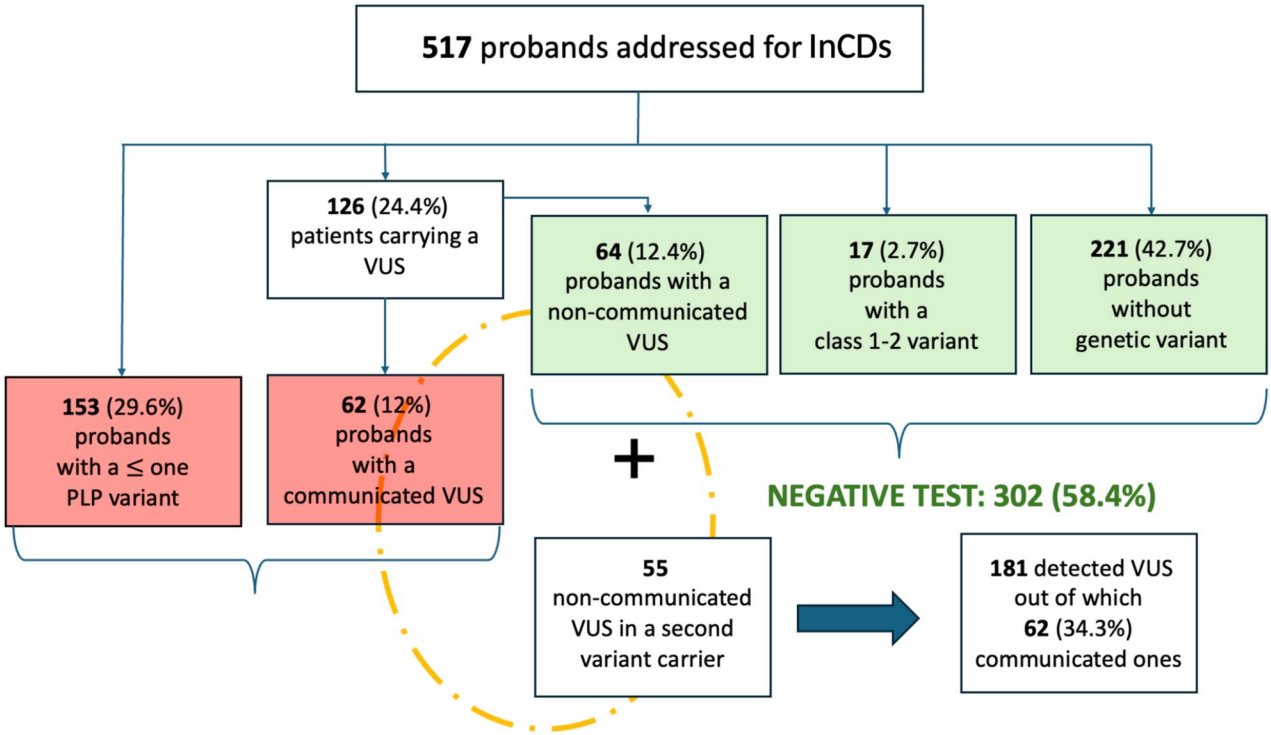

Fig. 1 Diagnostic yield of the genetic test. Genetic test results are displayed: the test was negative (302 pts) either for the absence of detected variant (221 pts) or due to the identification of non-diagnostic actionable variants (81 pts including 17 pts with class 1–2 variants and 64 pts with non-communicated VUS). Variant detection was communicated to the remaining 215 pts (including 153 pts with class 4–5 variants and 62 pts with communicated VUS). Regarding the 126 isolated VUS detected, 55 additional non-communicated VUS were observed in pts already carrying another variant. Among the 181 VUS globally detected, the result was communicated to pts only in 62 cases. *InCD* inherited cardiac disease, *PLP* pathogenic or likely pathogenic, *pts* patients, *VUS* variant of uncertain significance

The remaining 30 variant reclassifications (43.5%) involved genes that currently lack definitive strong evidence of association with either channelopathies or cardiomyopathies. Moreover, when evaluating the reclassification type across gene groups, we observed that most class changes were represented by downgrades (from PLP to VUS or from VUS to BLB) among cardiomyopathy genes (9/10, 90%) and genes without strong association with either channelopathies or cardiomyopathies (29/30, 97%). In contrast, among channelopathy genes, 34% (10/29) of reclassifications were represented by upgrades (from VUS to PLP).

3.2 Predictor Benchmarking

We evaluated the predictive performance of MutScore on the specific subsets of InCD variants and compared it with the discriminatory power of four other major tools (Polyphen-2 [10], CADD [11], Revel [12] and Alpha-missense [13]) commonly used in clinical practice. Table 2a (see the electronic supplementary material) summarizes prediction scores from each tool for every variant. MutScore had the highest AUCs, with a discriminatory power significantly greater than that of Polyphen-2 (0.88 [CI 95% 0.65–0.84] vs 0.72 [CI 95% 0.65–0.79], $p = 0.0001$) and CADD (0.89 [CI 95% 0.85–0.94] vs 0.7 [CI 95% 0.63–0.76], $p < 0.0001$) and with a trend towards a greater performance than Alpha-missense (0.89 [CI 95%0.85–0.94] vs 0.83 [CI 95% 0.77–0.88], $p = 0.06$).

Compared to Revel, the second-best predictor, MutScore exhibited a comparable predictive performance (0.89 [CI 95% 0.85–0.94] vs 0.88 [CI 95% 0.84–0.92], $p =$ not significant [NS]), but had a better sensitivity (73% vs 57%) at the maximum tolerated false-positive rate of 10% (Fig. 3).

We further focused our analysis on the 152 missense variants interpreted as BLB (59) or PLP (93) at the final updated manual classification complying with the ACMG recommendations. We evaluated the accuracy of MutScore and Revel (the second-best predictor) as well as their reliability for correctly classifying variants by applying tool-specific thresholds for BLB and PLP categorization. The analysis of performance metrics was performed first excluding variants with inconclusive tool-specific results (i.e., with a score above the tool-specific threshold identifying BLB variants and below the PLP one) and secondly considering variants

Table 1 Gene distribution of detected variants

Gene	*N* (%)	Gene	*N* (%)
SCN5 A	45 (17.9)	*DES*	2 (0.8)
RYR2	27 (10.7)	*GJA1*	2 (0.8)
KCNH2	26 (10.3)	*KCNE1*	2 (0.8)
KCNQ1	25 (9.9)	*NKX2-5*	2 (0.8)
DSP	12 (4.8)	*PKP2*	2 (0.8)
AKAP9	10 (4)	*RBM20*	2 (0.8)
CACNA1 C	9 (3.6)	*SCN4B*	2 (0.8)
ANK2	7 (2.8)	*TNNI3*	1 (0.4)
HCN4	7 (2.8)	*SNTA1*	1 (0.4)
SCN10 A	7 (2.8)	*SLMAP*	1 (0.4)
LMNA	6 (2.4)	*RANGRF*	1 (0.4)
DSC2	6 (2.4)	*PRKAG2*	1 (0.4)
DSG2	4 (1.6)	*NPPA*	1 (0.4)
TRPM4	4 (1.6)	*MYLK2*	1 (0.4)
TRPM7	4 (1.6)	*KCNE5*	1 (0.4)
FLNC	4 (1.6)	*KCNE3*	1 (0.4)
TMEM43	3 (1.2)	*KCNE2*	1 (0.4)
TRDN	3 (1.2)	*CTNNA3*	1 (0.4)
KCNJ8	3 (1.2)	*CDH2*	1 (0.4)
KCNJ5	3 (1.2)	*CAV3*	1 (0.4)
DPP6	3 (1.2)	*CASQ2*	1 (0.4)
CACNB2	2 (0.8)	*ABCC9*	1 (0.4)
TJP1	2 (0.8)	*CACNA2D1*	1 (0.4)

with inconclusive scores as false-positive or false-negative results.

After excluding variants with inconclusive scores, both algorithms achieved optimal sensitivities (0.97 for MutScore and 1 for the Revel), but Revel presented with a significantly lower specificity (0.36 vs 0.83 for MutScore). This latter also presented a significantly higher false-positive rate (0.64 vs 0.17 for MutScore) (Fig. 4, panel A).

When considering the intention-to-diagnose analysis, in order to evaluate the real diagnostic accuracy in clinical practice (after reallocating inconclusive results as false positive or false negative), both algorithms achieved good sensitivities (0.83 for MutScore and 0.92 for Revel), but their specificities markedly dropped, with a particularly poor value for Revel (0.14 vs 0.32 for MutScore). This latter also presented a higher false-positive rate (0.86 vs 0.68 for MutScore) (Fig. 4, panel B).

The concordance rate between the two predictors was 4.6% (7/152), 17.1% (26/152) and 50% (76/152) for BLB variants, VUS and PLP variants, respectively, with an overall concordance rate for 71.7% of variants. We further evaluated the true concordance rate (correct classification by concordant predictors), the false concordance rate (incorrect classification by concordant predictors) and the discordance rate (discordant calculator predictions). Among variants for which calculator predictions were concordant, we observed that 100% (7/7) and 97.4% (74/76) of variants were correctly classified as BLB and PLP, respectively, by concordant predictors. Of note, two out of the 76 variants interpreted as PLP by both predictors corresponded to BLB variants at the last updated classification.

Globally, we calculated a discordance rate of 28.3% (43/152). Interestingly, predictors were either discordant or

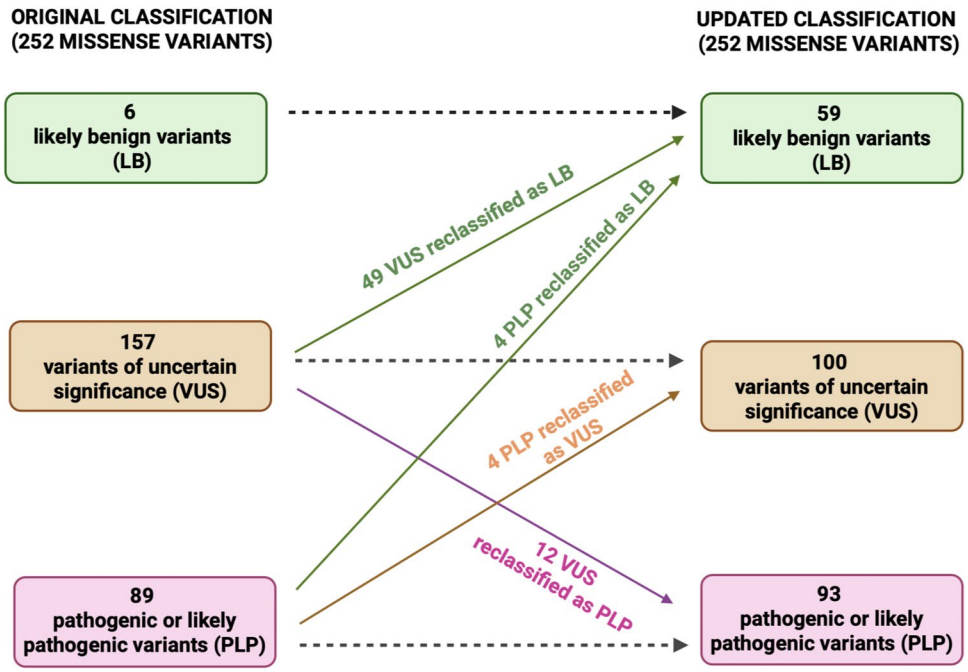

Fig. 2 Variant reclassifications occurred over the 9-year follow-up period

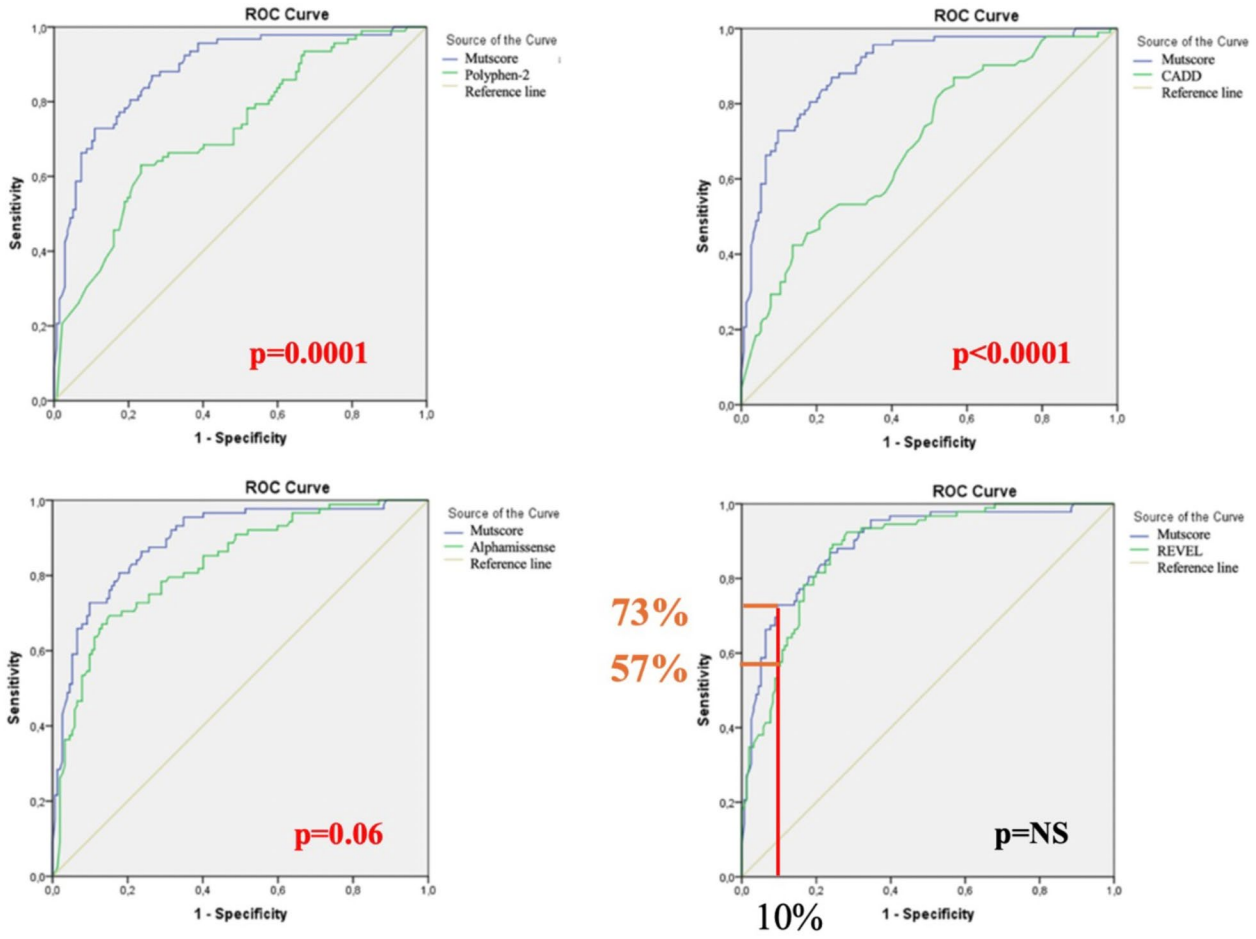

Fig. 3 ROC curves of tested predictors. *NS* not significant, *ROC* receiver operating characteristic

falsely concordant for 34.2% of BLB variants against 12.5% of PLP variants.

Finally, as shown in Fig. 4 (Table 1 in both panels, raw "MutScore + Revel"), by applying the consensus-based approach suggested by the ACMG recommendations when dealing with multiple tools, we calculated a sensitivity ranging from 100% to 96% and a specificity from 78% to 22% for the concordance of MutScore and Revel, depending on whether variants with inconclusive score were excluded or reallocated. However, a major limitation of this approach lies in the difficult management of discordant cases (i.e., variants for which prediction tools are discordant), which encompassed 28.3% of variants.

3.3 Potential VUS Reclassification

We investigated the potential of MutScore and Revel to disambiguate the 157 variants classified as VUS at time of the original classification. By applying tool-specific thresholds for BLB and PLP categorization, MutScore supported the potential reclassification of 44.6% (70/157) of VUS; 18.5% (29/157) were suggested to be BLB variants and 26.1% (41/157) PLP variants. Revel suggested reclassification for 42.7% of VUS, including 7% (11/157) redirected as BLB variants and 35.7% (56/157) potentially reclassifiable as PLP variants (Fig. 5).

In addition, 63.3% (7/11) of variants reclassifiable by Revel into BLB were downgraded to BLB variants, while 19.6% (11/56) of variants potentially reclassifiable by Revel as PLP, were upgraded to PLP variants at the last updated classification. Considering MutScore, 55.2% (16/29) of VUS reclassifiable as BLB were currently downgraded to BLB variants, while 19.5% (8/41) of VUS potentially reclassifiable in PLP were confirmed as PLP variants at the last updated classification.

3.4 Channelopathies Versus Cardiomyopathies

We also investigated whether MutScore and the best predictor among the four other tested provide consistent performance between channelopathy- and cardiomyopathy-related variants. We focused our analysis on the real-life dataset

A

	True positive	True negative	False positive	False negative	False positive rate	False negative rate	Sensitivity	Specificity	Accuracy	LR+	LR-
Mutscore	77	19	4	2	0.17	0.03	0.97	0.83	0.94	5.7	0.04
Revel	86	8	14	0	0.64	0	1	0.36	0.87	1.5	0
Mutscore + Revel	74	7	2	0	0.22	0	1	0.78	0.98	4.5	0

	BLB	PLP	Total
Mutscore BLB	19 (TN)	2 (FN)	21
Mutscore PLP	4 (FP)	77 (TP)	81
Total	23	79	102

	BLB	PLP	Total
Revel BLB	8 (TN)	0 (FN)	8
Revel PLP	14 (FP)	86 (TP)	100
Total	22	86	108

B

	True positive	True negative	False positive	False negative	False positive rate	False negative rate	Sensitivity	Specificity	Accuracy	LR+	LR-
Mutscore	77	19	40	16	0.68	0.17	0.83	0.32	0.63	1.2	0.53
Revel	86	8	51	7	0.86	0.08	0.92	0.14	0.62	1.1	0.57
Mutscore + Revel	74	7	25	3	0.78	0.04	0.96	0.22	0.74	1.2	0.18

	BLB	PLP	Total
Mutscore BLB	19 (TN)	2 (FN)	21
Mutscore inconclusive	36	14	50
Mutscore PLP	4 (FP)	77 (TP)	81
Total	59	93	152

	BLB	PLP	Total
Revel BLB	8 (TN)	0 (FN)	8
Revel inconclusive	37	7	44
REVEL PLP	14 (FP)	86 (TP)	100
Total	59	93	152

Fig. 4 **a** Variant classification according to predictor-specific thresholds and tool-specific contingency tables (after exclusion of variants with inconclusive results). **b** Variant classification according to predictor-specific thresholds and tool-specific contingency tables (after reallocating variants with inconclusive results). *BLB* benign/likely benign variants (according to final updated classification), *FN* false negative, *FP* false positive, *LR−* negative likelihood ratio, *LR+* positive likelihood ratio, *MutScore BPB* variants classified as BLB according to MutScore, *MutScore PLP* variants classified as PLP according to MutScore, *PLP* pathogenic/likely pathogenic variants (according to final updated classification), *Revel BPB* variants classified as BLB according to Revel, *Revel PLP* variants classified as PLP according to Revel, *TN* true negative, *TP* true positive

variants located on genes associated with definite/strong evidence of causality to channelopathies or cardiomyopathies according to current recommendations [8]. On this basis, we reported 140 variants annotated on channelopathy genes (CHANVs) and 42 variants annotated on cardiomyopathy genes (MYOVs). While both MutScore and Revel performed similarly on CHANVs (AUC of 0.88 [ICI 95% 0.82–0.94] vs 0.87 [CI 95% 0.81–0.93], p = NS), only MutScore achieved a significant better prediction performance for MYOVs (AUC of 0.77 [CI 95% 0.59–0.95], $p = 0.039$ vs Revel AUC of 0.61 [CI 95% 0.39–0.82], $p = 0.408$). Of note, however, when comparing MutScore AUCs in the two variant groups, we observed that it performed better on CHANVs than on MYOVs (AUC of 0.88 vs AUC of 0.77).

4 Discussion

The present study validates the excellent performance of MutScore, a new prediction tool for missense variants, in a real-life dataset of variants associated with the rare subset of InCDs. Our study first corroborates the consistent prediction accuracy when applying this new tool on the variant subgroup of InCD, poorly represented in training datasets.

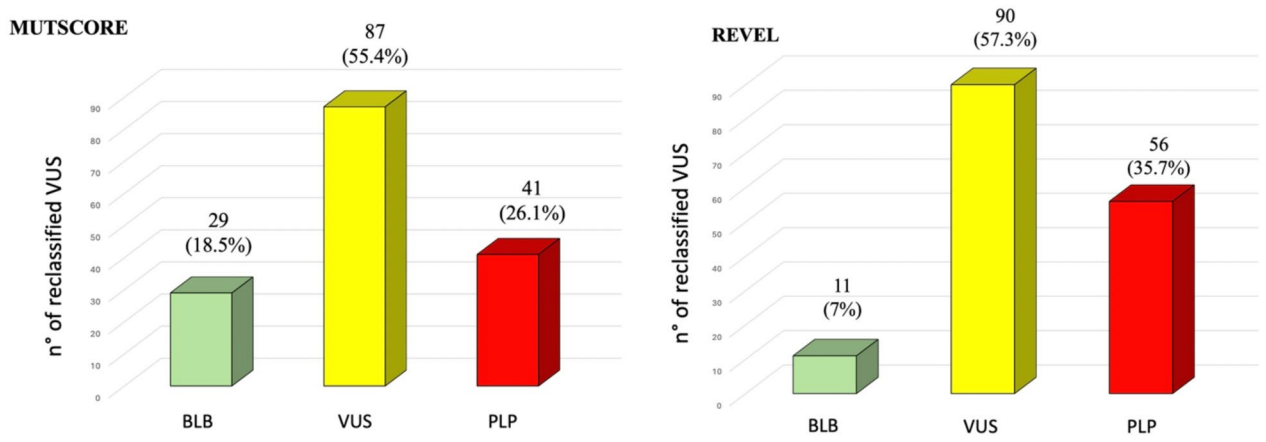

Fig. 5 VUS reclassification according to tool-specific thresholds. *BLB* benign/likely benign variants, *PLP* pathogenic/likely pathogenic variants, *VUS* variants of uncertain significance

4.1 Key Messages from the Global InCD Variant Dataset

Over a 9-year experience in a cohort of more than 500 probands with InCDs, we reported an overall diagnostic yield (due to the detection of clinically actionable PLP variants) in 29.6% of patients. Our results align with previous studies describing a diagnostic yield around 20–30% [20–23], when applying NGS technologies to channelopathy and cardiomyopathy panels. Focusing on original classification, we detected one or more VUS among 35% of probands, including 24.4% carrying an isolated VUS, 2.1% carrying a VUS and a second non-clinically actionable variant and 8.5% carrying a VUS and a second communicated variants. Globally, 26.5% of probands presented with one or more detected VUS in the absence of any other causally explaining variants. Considering the literature, VUS prevalence from other cohorts of InCD patients is quite variable [20–26]. Dellefave-Castillo et al. [20] described VUS detection in 51.2% of their InCD patients. Conversely, Young et al. [24] recently pointed out VUS detection in 18.2% of their InCD patients, while other works reported a VUS prevalence ranging from 21% [25] to 69.4% [26] when considering only cohorts of patients suffering from inherited channelopathies. Such results clearly outline a major discrepancy among studies, probably reflecting significant differences among gene panels and number of tested genes. As demonstrated by van Lint et al. [22], the yield of VUS significantly increased with the increasing number of genes included in the panels. Since laboratory protocols and gene panels may vary among centers as well as according to patients' phenotypes, the different panel choice may lead to inconsistent findings, making challenging comparisons and hindering the ability to draw definite conclusions on VUS prevalence among InCD patients. Moreover, although variant interpretation should normally comply with the ACMG recommendations, a certain degree of methodological heterogeneity in applying classification criteria may further contribute to result inconsistency. Interestingly, among the 181 VUS carriers of our cohort, genetic results were disclosed to patients only in 34.3% of cases. In our experience, indeed, the decision about VUS disclosure was not systematic but guided by specific criteria. VUS were reported to patients only when considered highly suspicious (i.e., "hot VUS") according to multiple supporting features—such as variant location in mutational hotspots, concordant deleterious predictions from multiple in silico tools, etc.—but lacking sufficient evidence for classification as likely pathogenic variants (i.e., due to the absence of functional studies or previous report in literature). Once such VUS were identified, the possibility to perform segregation analysis within the family represented the crucial criterion for disclosure, aiming at providing further evidence to support reclassification. This targeted and rationale-driven approach was intended to balance clinical engagement and responsibility, while avoiding unnecessary confusion among patients and family members. Indeed, recent literature analyzing the ethical and psychosocial implications associated with disclosure of inconclusive findings has increasingly drawn attention to the state of confusion and anxiety raised from the routine and indiscriminate communication of VUS to patients [27–29]. In addition, a further major challenge in variant interpretation and subsequent disclosure is the co-detection of multiple variants within the same patient, especially in the case of concomitant VUS. While each single variant may not fulfil pathogenicity criteria, it is conceivable that their synergistic effect might explain the patient's pathological phenotype. Further studies will be necessary to address this complex issue and to clarify the clinical significance as well as the implications for disclosure and counseling.

Noteworthy, about VUS reinterpretation, we documented that 38.9% of VUS initially detected have been successfully disambiguated on a 9-year follow-up. Interestingly, 80.4% of VUS were reclassified into BLB variants. Such results are in line with previous studies [24, 25, 30, 31] showing that VUS more often change to BLB variants essentially due to progressive improvement in knowledge about global population data [25, 26]. The increasing awareness of minor allele frequencies in the global population represents, indeed, a major population criterion to classify BLB variants. More broadly, nearly 30% of missense variants underwent a change in classification over the 9-year follow-up. Among these, reclassification was communicated to patients in one tier of cases, given the significant impact on clinical management (i.e., VUS upgraded to PLP and PLP variants downgraded to VUS or to BLB). Our results further highlight the pivotal importance of periodic reanalysis of variants, incorporating newly available scientific evidence, due to the potential implications for patient management, as attested by a 4.8% increase in the molecular diagnosis of InCDs over the 9-year period. Our findings are in line with previous studies analyzing the reclassification of rare variants associated with InCDs, despite the yield of reassessment varying depending on the duration of follow-up. Considering inherited channelopathies, Smith et al. [32] reported a 3% reclassification rate after 1 year, while Sarquella-Brugada et al. [26] demonstrated that nearly 20% of variants were reclassified after 5 years. Similarly, Cherny et al. [33] observed in a cohort of patients with InCDs that 22% of variants changed classification after 11 years, significantly altering clinical management in about 10% of cases. However, the optimal timing for variant re-evaluation remains a matter of debate, since current ACMG guidelines do not indicate a recommended interval for variant reassessment.

Finally, we detected 252 missense variants in our cohort, distributed on 46 different genes. The most represented ones were *SCN5 A* (17.9%), *RYR2* (10.7%), *KCNH2* (10.3%) and *KCNQ1* (9.9%). Globally, 55.2% of missense variants were located on genes associated with definite/strong evidence of causality to channelopathies. Several factors may contribute to these results, which were unexpected if we consider the higher prevalence of cardiomyopathies, such as hypertrophic or dilated cardiomyopathy, as compared to channelopathies. Despite the absence of available data on genetic test indications, our results may reflect the long-standing and well-established expertise of our center in channelopathies, leading to the recruitment of patients suffering from channelopathies and, hence, high representation of related gene variants. In addition, the significant prevalence of truncating variants on cardiomyopathy genes (i.e., *MYBC3*, *TTN*) might be an additional explanation.

4.2 Key Messages from Predictor Benchmarking

Herein, we found that MutScore exhibited the best predictive performance with the highest AUC among all tested algorithms. Such results are in line with the results of Quinodoz et al. [15], who validated MutScore on three different independent datasets. They showed that MutScore markedly outperformed traditional tested tools (including Polyphen, CADD and Revel), among which Revel already emerged as the second-best predictor. Similarly, Brock et al. [34] found that MutScore achieved the highest accuracy among 39 classifier tools for autosomal dominant variants related to inherited retinal diseases. Compared to previous works, our study not only further validated MutScore in a new real-life dataset but provided important evidence that MutScore maintained high and consistent prediction accuracy when applied to gene/variant types, which are underrepresented in training datasets, such as genes implicated in InCDs. By demonstrating for the first time the optimal predictive performance of MutScore in the field of cardiogenetics, our study supported its applicability for InCDs.

We also focused our analysis on the 152 missense variants classified as BLB or PLP at the last updated manual classification and applied the tool-specific thresholds for BLB and PLP categorization. Despite both achieving good sensitivities, MutScore provided a better specificity and a significantly lower false-positive rate. These data indicate that MutScore was globally more proficient in identifying BLB variants (i.e., true negatives), ensuring that fewer BLB variants were mistakenly classified as false-positive PLP variants. The analysis of concordance/discordance rate similarly confirmed the suboptimal performance of Revel in discriminating BLB variants since among variants classified as BLB by MutScore and as VUS by Revel, 92.3% corresponded to BLB variants at the last updated classification. Analogously, among variants interpreted as VUS by MutScore and as PLP by Revel, 52.2% were finally classified as BLB at the manual updated classification.

Interestingly, both predictors were either discordant or falsely concordant for 34.2% of BLB variants against 12.5% of PLP variants, indicating a general tendence of prediction tools to less accurately predict variant benignity. Of note, beyond benchmarking prediction tools, our findings clearly demonstrated that computational data—regardless of the specific in silico tool used—should not be considered as stand-alone criteria for variant interpretation. Conversely, they should always be integrated in the multi-layered workflow recommended by current ACMG standards including, among others, population data, segregation analysis and evidence from literature and disease databases.

We explored the potential of MutScore and Revel to reassess and disambiguate 157 variants classified as VUS at the original classification by applying tool-specific

thresholds for BLB and PLP categorization. Despite a similar rate of potentially reclassifiable VUS (44.6% for MutScore and 42.7% for Revel), MutScore seemed to have a greater ability in retrieving BLB variants as demonstrated by the greater rate of VUS categorized as BLB variants by MutScore (18.5%) as compared to Revel (7%). Moreover, such MutScore-driven reclassification seems to be already corroborated by current evidence. Indeed, the 55.2% of VUS reclassifiable as BLB variants have been currently downgraded to BLB variants according to the last updated classification. Of note, the work by Quinodoz et al. [15] already pointed out that MutScore had an edge in redirecting VUS toward BLB variants compared to other algorithms. We reasoned that this MutScore prerogative might be attributed to some unique features of its algorithm, which differs from those of other predictors, despite a partial degree of circularity and information redundancy, as between MutScore and Revel. Indeed, despite both sharing several features (i.e., phastCons, phyloP, GERP++RS, LRT, SIFT and PROVEAN), MutScore essentially differs for two novel scores. These latter include the positional score and the amino acid change score, integrating the information derived from the within-gene positional clustering of genetic variants. We may thus speculate that these topographic data contribute to a better variant classification (in relation to their localization inside or outside the identified clusters) and that such information is discriminating particularly for BLB variants.

In summary, although Revel currently represents one of the most effective predictors, which has been extensively evaluated in large-scale benchmarking studies [5, 12, 35], we demonstrated that MutScore globally outperformed Revel. Indeed, despite comparable sensitivities, MutScore provided an additional value due to a better specificity and a significantly lower false-positive rate in our real-world cohort. This peculiar performance advantage, which minimizes the mistakes of classifying BLB variants into false-positive PLP variants, may be particularly relevant in the clinical setting to avoid the unnecessary anxiety and the practical interventions associated with the over-allocation of BLB variants as PLP variants. Moreover, another advantage of MutScore is represented by its user-friendly application in clinical practice, due to the availability of two-sided thresholds for defining BLB and PLP variants. These values were originally defined in the work by Quinodoz et al. [15] and correspond to MutScore cutoffs that correctly identify 95% of BLB and PLP variants, respectively, in the original dataset used to train the algorithm. These thresholds, by reflecting a 95% sensitivity for both BLB and PLP variant detection, identify two high confidence regions for the assessment of true-positive (PLP) and true-negative (BLB) results, while leaving an intermediate region of uncertainty for VUS. The application of such cutoffs in our study allowed us to evaluate MutScore predictive performance and to further corroborate their validity and clinical relevance in a real-world dataset.

4.3 Key Messages from the Comparison Between Channelopathy and Cardiomyopathy-Related Variants

We found that both MutScore and Revel performed similarly on CHANVs. However, only MutScore achieved a significant prediction performance on MYOVs. Despite the smaller MYOV sample size, MutScore-specific topographic scores may explain such results, by providing a more nuanced and accurate variant assessment as compared to features used by other algorithms.

We also observed that MutScore seemed to potentially better perform on CHANVs than on MYOVs, yielding an AUC of 0.88 versus 0.77, respectively, although this finding should be interpreted with caution due to the limited size of the MYOV sample. A recent study [36] demonstrated that PLP missense variants in cardiovascular genes displayed a higher clustering (i.e., an increased regional density within specific gene regions) compared to truncating variants. This finding may be explained by the fact that truncating variants essentially cause haploinsufficiency/loss of function regardless of their transcript location, while missense variants more often induce dominant-negative mechanisms which depend on the alteration of specific functional domains [36]. This aspect is particularly relevant for channelopathies, where dominant-negative effects are well-established mechanisms of pathogenicity. Accordingly, missense variants often account for the majority of PLP variants (i.e., 86–92% of *RYR2* PLP variants [37] or up to 61% of *KCNQ1* PLP variants in long QT syndrome [38]). Conversely, at least for some major cardiomyopathy genes such as *MYBPC3* or *TTN*, the majority of variants interpreted as PLP are truncating, acting via haploinsufficiency [39, 40]. For instance, truncating variants represent up to 90% of PLP variants in *MYBPC3* [39] as well as the large majority of PLP variants in *TTN* [41], whereas missense variants more often do not fulfil pathogenicity criteria due to insufficient supporting evidence. Based on these observations, we may expect a better performance of MutScore in gene contexts where PLP missense variants present domain-specific distribution, as is frequently the case in channelopathy genes. This enhanced accuracy may be potentially explained by MutScore incorporation of topographic features that derive from the known location of variants within the gene/protein structure.

4.4 Limitations

The major limitation of our study is represented by its retrospective approach, which might introduce variability in data quality or completeness and hamper the rigorous

collection of genetic variants. This is particularly true for BLB variants that might not have been fully captured or accurately reported due to the related absence of clinical actionability. The results of our real-life dataset should be interpreted with caution, accounting for the fact that some genetic variants (i.e., BLB) may be underestimated. Secondly, we do mention that different panel choices tailored on patients' phenotypes may have impacted genetic analyses on the prevalence of detected VUS. Another major issue is the absence of a gold-standard for correct variant interpretation. We assumed the last updated classification complying with ACMG recommendations as the more reliable approximation. This approach does not rule out the uncertainty that persists for some dataset variants still classified as VUS and does not completely eliminate the risk of evolving classification (according to new available scientific evidence). Fourthly, considering that according to Bayes theorem, the positive/negative predictive values depend on the prevalence of the disease, we did not calculate the positive/negative predictive values adjusted for disease prevalence. However, to account for this limitation, we calculated the likelihood ratios that are independent from disease prevalence. Finally, the limited size of the MYOV sample may make it underpowered to detect existing significant differences.

5 Conclusions

Our study supports the excellent performance of MutScore in a real-life dataset of variants associated with the rare subset of InCDs. Beyond corroborating the consistent prediction accuracy on this variant subgroup poorly represented in training datasets, we found that MutScore was globally more proficient at classifying variants compared to the other four predictors commonly used in clinical practice (Polyphen-2 [10], CADD [11], Revel [12] and Alpha-missense [13]). As compared to Revel, the second-best predictor, MutScore exhibited a better sensitivity at the maximum tolerated false-positive rate of 10%, a better specificity and a markedly lower false-positive rate, supporting a more nuanced and accurate assessment especially for BLB variants. We also demonstrated that MutScore seemed to perform better on CHANVs than on MYOVs. MutScore represents a promising tool in cardiogenetics that can contribute to the translation of personalized medicine into practice.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40291-025-00784-8.

Declarations

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on request.

Funding statement Open access funding provided by University of Lausanne. The authors have nothing to declare regarding the preparation of this manuscript.

Author contributions APP: substantial contributions to the conception and design of the work, to the analysis and interpretation of data and to drafting and substantively revising the work. VF: substantial contributions to the acquisition, analysis and interpretation of data and to substantively revising the work. ES: contribution to substantively revising the work. CM: contribution to substantively revising the work. AB: contribution to substantively revising the work. AM: contribution to substantively revising the work. VA: contribution to substantively revising the work. GV: contribution to substantively revising the work. EP: contribution to substantively revising the work. AL: contribution to substantively revising the work. ID: contribution to substantively revising the work. FE: contribution to the conception and design of the work, to the analysis and interpretation of data and to substantively revising the work.

Ethics declaration All patients or their legal representatives provided signed, informed consent to allow the collection of personal clinical and genetic data and their use for research purposes. The data were collected in compliance with the national French data protection regulations (*Commission Nationale de l'Informatique et des Libertés* [CNIL]). Due to the design of the study being based on routine clinical practice, approval of the study protocol by an institutional review board was not necessary. Conversely, according to French legislation, the study protocol was approved by the appropriate hospital committee and complied with the Declaration of Helsinki 1975 and with its further amendments.

Conflict of interest The authors (APP, VF, ES, CM, AB, AM, VA, GV, EP, AL, ID, FE) have no conflicts of interest to declare.

Consent (participation and publication) Not applicable.

Code availability MutScore is available from https://iob-genetic.shinyapps.io/mutscore/ or from https://iob-genetic.shinyapps.io/mutscore-batch, and it can be used under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. For commercial uses of MutScore, please contact Carlo Rivolta (carlo.rivolta@iob.ch).

References

1. Richards S, Aziz N, Bale S, et al., ACMG Laboratory Quality Assurance Committee. 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(5):405-24.
2. Boycott KM, Vanstone MR, Bulman DE et al. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. Nat Rev Genet. 2013.
3. Schwartz PJ, Ackerman MJ, Antzelevitch C, et al. Inherited cardiac arrhythmias. Nat Rev. 2020;6:58.
4. Cerrone M, Remme CA, Tadros R, Bezzina CR, Delmar M. Beyond the one gene-one disease paradigm: complex genetics and pleiotropy in inheritable cardiac disorders. Circulation. 2019;140(7):595–610.
5. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, Cannon-Albright LA, Teerlink CC, Stanford JL, Isaacs WB, Xu J, Cooney KA, Lange EM, Schleutker J, Carpten JD, Powell IJ, Cussenot O, Cancel-Tassin G, Giles GG, MacInnis RJ, Maier C, Hsieh CL, Wiklund F, Catalona WJ, Foulkes WD, Mandal D, Eeles RA, Kote-Jarai Z, Bustamante CD, Schaid DJ, Hastie T, Ostrander EA, Bailey-Wilson JE, Radivojac P, Thibodeau SN, Whittemore AS, Sieh W. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet. 2016;99(4):877–85.
6. Gunning AC, Fryer V, Fasham J, et al. Assessing performance of pathogenicity predictors using clinically relevant variant datasets. J Med Genet. 2021;58(8):547–55.
7. Livesey BJ, Marsh JA. Updated benchmarking of variant effect predictors using deep mutational scanning. Mol Syst Biol. 2023;19(8):e11474. https://doi.org/10.1525/msb.202211474. (**Epub 2023 Jun 13**).
8. Wilde AAM, Semsarian C, Márquez MF, et al. European Heart Rhythm Association (EHRA)/Heart Rhythm Society(HRS)/Asia Pacific Heart Rhythm Society (APHRS)/Latin American HeartRhythm Society (LAHRS) Expert Consensus Statement on the state of genetic testing for cardiac diseases. Europace. 2022;24(8):1307–67.
9. Zeppenfeld K, Tfelt-Hansen J, de Riva M, Winkel BG, Behr ER, Blom NA, Charron P, Corrado D, Dagres N, de Chillou C, Eckardt L, Friede T, Haugaa KH, Hocini M, Lambiase PD, Marijon E, Merino JL, Peichl P, Priori SG, Reichlin T, Schulz-Menger J, Sticherling C, Tzeis S, Verstrael A, Volterrani M. 2022 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. Eur Heart J. 2022;43(40):3997–4126.
10. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248–9.
11. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310–5.
12. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet. 2016;99(4):877–85. https://doi.org/10.1016/j.ajhg.2016.08.016.
13. Cheng J, Novati G, Pan J, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. Science. 2023;381(6664):eadg7492.
14. Baux D, Van Goethem C, Ardouin O, Guignard T, Bergougnoux A, Koenig M, Roux AF. MobiDetails: online DNA variants interpretation. Eur J Hum Genet. 2021;29(2):356–60.
15. Quinodoz M, Peter VG, Cisarova K, et al. Analysis of missense variants in the human genome reveals widespread gene-specific clustering and improves prediction of pathogenicity. Am J Hum Genet. 2022;109(3):457–70.
16. Schuetz GM, Schlattmann P, Dewey M. Use of 3x2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: meta-analytical evaluation of coronary CT angiography studies. BMJ. 2012;24(345): e6717.
17. Mannhart D, Lefebvre B, Gardella C, Henry C, Serban T, Knecht S, Kühne M, Sticherling C, Badertscher P. Clinical validation of an artificial intelligence algorithm offering cross-platform detection of atrial fibrillation using smart device electrocardiograms. Arch Cardiovasc Dis. 2023;116(5):249–57.
18. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology. 1983;148(3):839–43.
19. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36
20. Dellefave-Castillo LM, Cirino AL, Callis TE, et al. Assessment of the diagnostic yield of combined cardiomyopathy and arrhythmia genetic testing. JAMA Cardiol. 2022;7(9):966–74.
21. Murphy J, Kirk CW, Lambert DM, et al. Diagnostic yield from cardiac gene testing for inherited cardiac conditions and re-evaluation of pre-ACMG variants of uncertain significance. Ir J Med Sci. 2024;193:1775.
22. van Lint FHM, Mook ORF, Alders M, Bikker H, LekanneDit-Deprez RH, Christiaans I. Large next-generation sequencing gene panels in genetic heart disease: yield of pathogenic variants and variants of unknown significance. Neth Heart J. 2019;27(6):304–9.
23. Janin A, Januel L, Cazeneuve C, et al. Molecular diagnosis of inherited cardiac diseases in the era of next-generation sequencing: a single center's experience over 5 years. Mol Diagn Ther. 2021;25(3):373–85.
24. Young WJ, Maung S, Ahmet S, Kirkby C, Ives C, Schilling RJ, Lowe M, Lambiase PD. The frequency of gene variant reclassification and its impact on clinical management in the inherited arrhythmia clinic. Heart Rhythm. 2024;21(6):903–10. https://doi.org/10.1016/j.hrthm.2024.01.008. (**Epub 2024 Jan 11**).
25. Bennett JS, Bernhardt M, McBride KL, Reshmi SC, Zmuda E, Kertesz NJ, Garg V, Fitzgerald-Butt S, Kamp AN. Reclassification of variants of uncertain significance in children with inherited arrhythmia syndromes is predicted by clinical factors. Pediatr Cardiol. 2019;40(8):1679–87.
26. Sarquella-Brugada G, Fernandez-Falgueras A, Cesar S, Arbelo E, Coll M, Perez-Serra A, Puigmulé M, Iglesias A, Alcalde M, Vallverdú-Prats M, Fiol V, Ferrer-Costa C, Del Olmo B, Picó F, Lopez L, García-Alvarez A, Jordà P, de Llano CT, Toro R, Grassi S, Oliva A, Brugada J, Brugada R, Campuzano O. Clinical impact of rare variants associated with inherited channelopathies: a 5-year update. Hum Genet. 2022;141(10):1579–89.
27. Werner-Lin A, Mccoyd JLM, Bernhardt BA. Actions and uncertainty: how prenatally diagnosed variants of uncertain significance become actionable. Hastings Cent Rep. 2019;49
28. Clift K, Macklin S, Halverson C, McCormick JB, Abu Dabrh AM, Hines SJ. Patients' views on variants of uncertain significance across indications. Community Genet. 2020;11(2):139–45.
29. Reuter C, Chun N, Pariani M, Hanson-Kahn A. Understanding variants of uncertain significance in the era of multigene panels: through the eyes of the patient. J Genet Couns. 2019;28(4):878–86.
30. Chen E, Facio FM, Aradhya KW, Rojahn S, Hatchell KE, Aguilar S, Ouyang K, Saitta S, Hanson-Kwan AK, Capurro NN, Takamine E, Jamuar SS, McKnight D, Johnson B, Aradhya S. Rates and

classification of variants of uncertain significance in hereditary disease genetic testing. JAMA Netw Open. 2023;6(10): e2339571.
31. Lahrouchi N, Raju H, Lodder EM, Papatheodorou E, Ware JS, Papadakis M, Tadros R, Cole D, Skinner JR, Crawford J, Love DR, Pua CJ, Soh BY, Bhalshankar JD, Govind R, Tfelt-Hansen J, Winkel BG, van der Werf C, Wijeyeratne YD, Mellor G, Till J, Cohen MC, Tome-Esteban M, Sharma S, Wilde AAM, Cook SA, Bezzina CR, Sheppard MN, Behr ER. Utility of post-mortem genetic testing in cases of sudden arrhythmic death syndrome. J Am Coll Cardiol. 2017;69(17):2134–45.
32. Smith ED, Radtke K, Rossi M, Shinde DN, Darabi S, El-Khechen D, Powis Z, Helbig K, Waller K, Grange DK, Tang S, Farwell Hagman KD. Classification of genes: standardized clinical validity assessment of gene-disease associations aids diagnostic exome analysis and reclassifications. Hum Mutat. 2017;38(5):600–8.
33. Cherny S, Olson R, Chiodo K, Balmert LC, Webster G. Changes in genetic variant results over time in pediatric cardiomyopathy and electrophysiology. J Genet Couns. 2021;30(1):229–36.
34. Brock DC, Wang M, Hussain HMJ, et al. Comparative analysis of in-silico tools in identifying pathogenic variants in dominant inherited retinal diseases. Hum Mol Genet. 2024;33(11):945–57.
35. Tian Y, Pesaran T, Chamberlin A, Fenwick RB, Li S, Gau C-L, Chao EC, Lu H-M, Black MH, Qian D. REVEL and BayesDel outperform other in silico meta-predictors for clinical variant classification. Sci Rep. 2019;9(1):12752.
36. Krysov VA, Wilson RH, Ten NS, et al. Regional variation in cardiovascular genes enables a tractable genome editing strategy. Circ Genom Precis Med. 2024;17(2):e004370.
37. Steinberg C, Roston TM, van der Werf C, Sanatani S, Chen SRW, Wilde AAM, Krahn AD. RYR2-ryanodinopathies: from calcium overload to calcium deficiency. Europace. 2023;25(6):euad156.
38. Novelli V, Faultless T, Cerrone M, Care M, Manzoni M, Bober SL, Adler A, De-Giorgio F, Spears D, Gollob MH. Enhancing the interpretation of genetic observations in KCNQ1 in unselected populations: relevance to secondary findings. Europace. 2023;25(11):euad317.
39. Helms AS, Thompson AD, Glazier AA, Hafeez N, Kabani S, Rodriguez J, Yob JM, Woolcock H, Mazzarotto F, Lakdawala NK, Wittekind SG, Pereira AC, Jacoby DL, Colan SD, Ashley EA, Saberi S, Ware JS, Ingles J, Semsarian C, Michels M, Olivotto I, Ho CY, Day SM. Spatial and functional distribution of MYBPC3 pathogenic variants and clinical outcomes in patients with hypertrophic cardiomyopathy. Circ Genom Precis Med. 2020;13(5):396–405.
40. Pugh TJ, Kelly MA, Gowrisankar S, Hynes E, Seidman MA, Baxter SM, Bowser M, Harrison B, Aaron D, Mahanta LM, Lakdawala NK, McDermott G, White ET, Rehm HL, Lebo M, Funke BH. The landscape of genetic variation in dilated cardiomyopathy as surveyed by clinical DNA sequencing. Genet Med. 2014;16(8):601–8.
41. Bourfiss M, van Vugt M, Alasiri AI, Ruijsink B, van Setten J, Schmidt AF, Dooijes D, Puyol-Antón E, Velthuis BK, van Tintelen JP, Te Riele ASJM, Baas AF, Asselbergs FW. Prevalence and disease expression of pathogenic and likely pathogenic variants associated with inherited cardiomyopathies in the general population. Circ Genom Precis Med. 2022;15(6): e003704.