Ruiz M, Ochoa JP, Migoyo-Bettoni C, de la Barrera J, Delrio-Lorenzo A, Fernández-Rojo MA, Martinez-Martin I, Alegre-Cebollada J, Lara-Pezzi E, Sanchez-Cabo F, Garcia-Pavia P. Performance of AlphaMissense and Other In Silico Predictors to Determine Pathogenicity of Missense Variants in Sarcomeric Genes. ***Circ Genom Precis Med***. 2025 Apr;18(2):e004922. doi: 10.1161/CIRCGEN.124.004922. Epub 2025 Mar 21. PMID: 40115995.

Performance of AlphaMissense and Other in-silico Predictors to Determine Pathogenicity of Missense Variants in Sarcomeric Genes

Running title: *Ruiz et al.; Evaluation of in-silico predictors*

Mario Ruiz, MSc[1]; Juan Pablo Ochoa, MD, PhD[1,2]; Candela Migoyo-Bettoni, MSc[1]; Jorge de la Barrera, MSc[1]; Alba Delrio-Lorenzo, PhD[1,3,4]; Manuel A. Fernández-Rojo, PhD[1]; Ines Martinez-Martin, MSc[1]; Jorge Alegre-Cebollada, PhD[1]; Enrique Lara-Pezzi, PhD[1,3]; Fatima Sanchez-Cabo, PhD[1]; Pablo Garcia-Pavia, MD, PhD[1,3-5]

[1]Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid; [2]Health in Code, A Coruña; [3]CIBERCV, Instituto de Salud Carlos III, Madrid; [3]Hospital Universitario Puerta de Hierro Majadahonda, IDIPHISA, Madrid; [4]Universidad Francisco de Vitoria (UFV), Pozuelo de Alarcón, Spain

Correspondence:

Pablo Garcia-Pavia, MD, PhD

Department of Cardiology

Hospital Universitario Puerta de Hierro

Manuel de Falla, 2

28222 Madrid, Spain.

Tel: (+ 34) 91 191 7297

Fax: (+34) 91 191 7718

E-mail: pablogpavia@yahoo.es

Nonstandard Abbreviations and Acronyms

VUS	Variant of Uncertain Significance
IRB	Institutional Review Board
P	Pathogenic
LP	Likely Pathogenic
B	Benign
LB	Likely Benign
ROC	Receiver Operating Characteristic
ACMG	American College of Medical Genomics
AUC	Area Under the Curve

Interpreting the pathogenicity of genetic variants remains a challenge. Even with advanced tools, many gene variants are still classified as Variants of Uncertain Significance (VUS), precluding their use in clinical practice. During the last 20 years a multitude of *in-silico* tools have been developed to predict the pathogenicity of variants and provide support for their interpretation.[1] AlphaMissense, a new *in-silico* predictor of every single missense change in the coding regions of the genome developed by Google DeepMind, combines population frequency information, an unsupervised language model to learn about amino acids distribution and context, and structural information through AlphaFold-derived modelling.[2] AlphaMissense distinguishing features include the use of a very accurate proprietary model of protein structure prediction based on deep learning and the absence of biases when selecting, filtering or curating the training dataset since it does not require human labeling.

AlphaMissense has shown good performance in addressing pathogenicity of missense variants present in ClinVar and has been used for establishing pathogenicity of variants in several conditions.[2-3] However, it has not been validated specifically in cardiovascular diseases.

We sought to evaluate AlphaMissense and other in-silico predictors in assessing the pathogenicity of missense variants in sarcomeric genes. The data that support the findings of this study are available from the corresponding author upon reasonable request. IRB approval was obtained.

For this purpose, we evaluated the performance of 12 predictors (AlphaMissense, REVEL, CADD, MutPred, PolyPhen2, MutationTaster, SIFT4G, FATHMM, DANN, PrimateAI, MetaLR and MetaRNN) when assessing missense variants identified in 8 sarcomeric genes (*MYBPC3*, *MYH7*, *TNNT2*, *TNNI3*, *TNNC1*, *ACTC1*, and *TPM1*) in 35,076 individuals with cardiovascular disorders sequenced at Health in Code (A Coruña, Spain).

The variants identified were categorized as Pathogenic/Likely Pathogenic (P, LP), VUS, or Benign/Likely benign (B/LB) following the recommendations of the ACMG.[4] Information available at Health in Code and in the literature regarding the number of individuals with the variant and their phenotype, familial history, and population frequency in gnomAD v2.1 were considered when establishing pathogenicity.

A total of 1,595 missense variants were identified. Upon review, 84 (5%) were considered B/LB, 1,181 (74%) VUS and 330 (21%) P/LP. P/LP and B/LB variants were used to evaluate the performance of the predictors.

Variants were annotated with the pre-computed AlphaMissense score file, from which both the score and its predicted pathogenicity (LB, ambiguous and LP) were extracted.[5] VEP (v.111) and dbNSFP (v4.3) were used for other predictors.

The pathogenicity values provided by each predictor were transformed into values ranging from 0 (benign) to 1 (pathogenic). Variants that could not been assessed by predictors were not considered for evaluating predictors' performance.

Performance and discriminant power of predictors were evaluated by the area under de curve (AUC) of the Receiver Operating Characteristic (ROC) curves. AlphaMissense had an acceptable discrimination performance (AUC=0.85). Nevertheless, 5 predictors (MutPred, MetaRNN, MetaLR, REVEL and PrimateAI) showed higher discriminant power (Figure 1B). Since the ratio of B/LB to P/LP variants was imbalanced, precision recall curves were obtained. The precision curve of AlphaMissense depicted good results (AP=0.96), but MutPred, MetaRNN, MetaLR and REVEL showed higher AP (Figure 1C). Notably, some predictors failed to predict (no prediction available) many variants in certain genes (i.e. MutPred and PolyPhen2 failed to predict 60% of *MYBPC3* variants, and PrimateAI 30% of *TPM1* variants).

To validate these findings and avoid evaluating pathogenicity using ACMG criteria that include computational evidence, same analysis was performed restricted to variants classified as P/LP or B/LB without using computational evidence, and to those variants with pathogenicity stablished by functional studies. In both subsets similar results were obtained.

The 4 predictors with the best performance were compared after excluding MutPred as it did not return a score for 124 (60%) variants (83.3% of B/LB variants).

MetaRNN was the best predictor in our dataset and the predictor that correctly classified more variants (Figure). Furthermore, MetaRNN could be used in all variants and gave very low values to variants classified as benign and very high to those classified as pathogenic. AlphaMissense was ranked third in variants correctly classified (284, 68.6%) despite it was the only predictor that correctly classified 9 variants (2.2%)(Figure 1D).

Based on the overall results we would conclude that the performance of AlphaMissense was inferior to that of MetaRNN and MetaLN and similar to that of REVEL.

Our study provides the first data about the performance of the recently described predictor AlphaMissense applied to cardiovascular genes and provides a complete benchmarking of *in-silico* predictors in this setting. Despite the acceptable performance of AlphaMissense, our results show that the performances of other predictors are superior. Our results would be useful in selecting the most appropriate *in-silico* tools to evaluate missense variants.

Sources of Funding: The CNIC is supported by the Instituto de Salud Carlos III (ISCIII), the Ministerio de Ciencia e Innovación (MCIN) and the Pro CNIC Foundation), and is a Severo Ochoa Center of Excellence (grant CEX2020-001041-S funded by MICIN/AEI/10.13039/501100011033). PGP and ELP are funded by the Pathfinder Cardiogenomics programme of the European Innovation Council of the European Union (DCM-NEXT project; project number: 101115416). J.A.-C. acknowledges funding from the Ministerio de Ciencia, Innovación y Universidades (MCIU, MICIU/AEI/10.13039/501100011033) through grant PID2020-120426GB-I00. FSC acknowledges funding from the Ministerio de Ciencia, Innovación y Universidades (MCIU, MICIU/AEI/10.13039/501100011033) and ERDF through grant PID2022-141527OB-I00.

Disclosures: None.

References:

1. Zadorozhny A, Smirnov A, Filimonov D, Lagunin A. Prediction of pathogenic single amino acid substitutions using molecular fragment descriptors. Bioinformatics. 2023;39:btad484.

2. Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., Pritzel, A., Wong, L. H., Zielinski, M., Sargeant, T., et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. Science 2023; 381(6664):eadg7492.

3. Chabane, K., Charlot, C., Gugenheim, D., Simonet, T., Armisen, D., Viailly, P. J., Codet de Boisse, G., Huet, S., Hayette, S., Alcazer, V., et al. Real life evaluation of alphamissense predictions in hematological malignancies. Leukemia 2023;38:420–3.

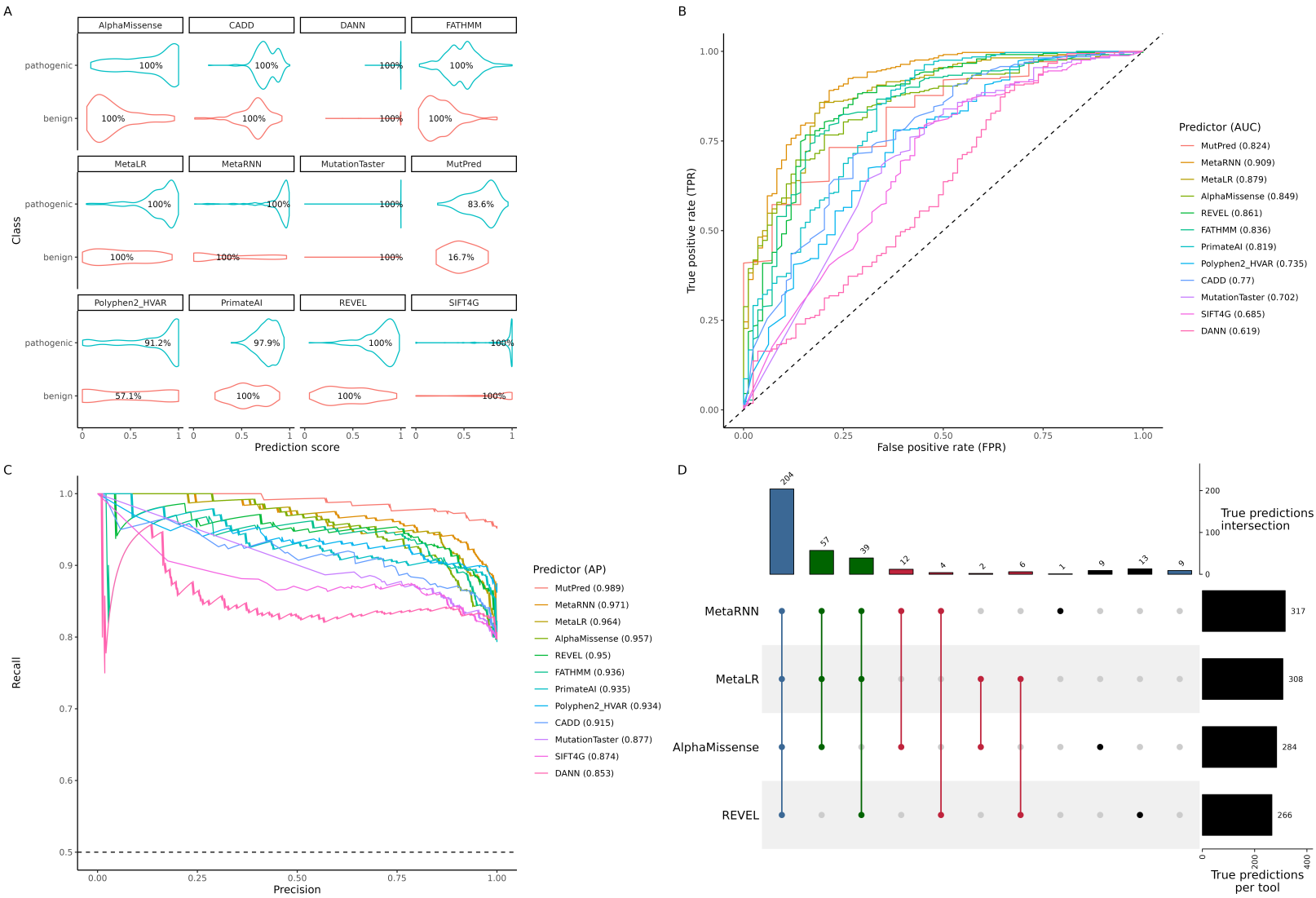
4. Miller, D. T., Lee, K., Abul-Husn, N. S., Amendola, L. M., Brothers, K., Chung, W. K., Gollob, M. H., Gordon, A. S., Harrison, S. M., Hershberger, R. E.,et al. ACMG SF v3.2 list for

reporting of secondary findings in clinical exome and genome sequencing. Genet Med 2023;25:100866.

5. https://console.cloud.google.com/storage/browser/dm_alphamissense

Figure Legend:

Figure 1. (A) Violin plot illustrating the distribution of pathogenicity scores according to classification of variants across predictors. The percentage of variants included in each class is displayed in the mean distribution value. (B) Receiver Operating Characteristic Curves for the 12 predictors (C) Precision-recall curves for the 12 predictors. (D) Upset plot displaying the intersection of correctly classified variants (true positives + true negatives) by the 4 predictors with better performance.


A
AlphaMissense
CADD
DANN
FATHMM
MetaLR
MetaRNN
MutationTaster
MutPred
Polyphen2_HVAR
PrimateAI
REVEL
SIFT4G
pathogenic
benign
Class
100%
83.6%
16.7%
91.2%
97.9%
57.1%
Prediction score
B
True positive rate (TPR)
False positive rate (FPR)
Predictor (AUC)
MutPred (0.824)
MetaRNN (0.909)
MetaLR (0.879)
AlphaMissense (0.849)
REVEL (0.861)
FATHMM (0.836)
PrimateAI (0.819)
Polyphen2_HVAR (0.735)
CADD (0.77)
MutationTaster (0.702)
SIFT4G (0.685)
DANN (0.619)
C
Recall
Precision
Predictor (AP)
MutPred (0.989)
MetaRNN (0.971)
MetaLR (0.964)
AlphaMissense (0.957)
REVEL (0.95)
FATHMM (0.936)
PrimateAI (0.935)
Polyphen2_HVAR (0.934)
CADD (0.915)
MutationTaster (0.877)
SIFT4G (0.874)
DANN (0.853)
D
204
57
39
12
4
2
6
1
9
13
9
True predictions intersection
MetaRNN
MetaLR
AlphaMissense
REVEL
317
308
284
266
True predictions per tool