Evaluation of Phenotype-driven Variant Prioritization Methods for Cardiogenetic Diagnosis in the Brazilian Population

Manuela C. P. Bonetto[1], Ed Carlos S. Silva[1], Juliana José[1]
Rogerio S. Rosa[1], Jose S. L. Patane[1,*]

[1] Computational Biology Laboratory - Laboratory of Genetics and Molecular Cardiology (InCor - HCFMUSP)
Av. Dr. Enéas Carvalho de Aguiar, 44 - São Paulo - SP - Brazil, CEP 05403-900

{manuela.bonetto, ed.silva, juliana.jose, rogerio.rosa, jose.patane}@hc.fm.usp.br

Abstract. *We conducted a systematic benchmarking of four phenotype-driven variant prioritization methods–hiPhive, PhenIX, LIRICAL, and xRare–using validated diagnoses from a cohort of 164 Brazilian patients with cardiogenetic conditions. Each method integrated HPO-coded phenotypes and Whole-Exome sequencing data to rank variants by predicted clinical relevance. Performance was quantified using top-K recall metrics. Exomiser-based approaches (hiPhive and PhenIX) demonstrated superior overall performance, with PhenIX showing greater consistency across disease subtypes and hiPhive excelling in top-1 rankings. LIRICAL achieved high recall in specific conditions, while xRare had lower accuracy overall. These results support the potential of ensemble strategies to enhance diagnostic precision, particularly important in resource-constrained healthcare settings like Brazil's public health system.*

1. Introduction

Inherited cardiopathies pose a critical public health challenge in Brazil, significantly affecting mortality and quality of life. In the public healthcare system (SUS), where demand for specialized services often exceeds capacity, expanding equitable access to genomic analyses is a practical necessity. The availability of open-access variant prioritization methods could reduce dependence on expensive commercial software.

Cardiogenetic conditions are among the most prevalent inherited disorders affecting the Brazilian population. The genetic heterogeneity and variable expressivity of these conditions make precise diagnosis challenging. Automated variant prioritization tools that integrate genomic and phenotypic data offer a promising solution to alleviate interpretation burdens.

This study evaluates four phenotype-driven open-source variant prioritization methods - Exomiser-hiPhive [Robinson et al. 2014], Exomiser-PhenIX [Smedley et al. 2015], LIRICAL [Robinson et al. 2020], and xRare [Li et al. 2019] - in a cohort of 164 Brazilian patients with genetically confirmed cardiac conditions. Using validated diagnoses as gold standard, we computed recall-at-top-K metrics to assess each tool's ability to retrieve the causal variant.

2. Methods

2.1. Patient Cohort and Data Preparation

We analyzed 164 Brazilian patients with validated cardiogenetic diagnoses from the Heart Institute (InCor). The cohort included patients with various cardiac conditions. Whole-exome sequencing was performed using standard protocols, with variants filtered for quality (depth $\geq$ 30x). Phenotypic information was extracted from medical records and converted to HPO terms [Köhler et al. 2021] by clinical geneticists.

2.2. Variant Prioritization Methods

We applied four computational tools (hiPhive, PhenIX, LIRICAL, xRare) to genomic (VCF) and phenotypic (HPO) data from 164 patients. Each tool processed preprocessed WES VCFs (quality-filtered, depth $\geq$ 30x). Phenotypic information was extracted from medical records and converted to HPO terms by clinical geneticists. The HPO curation process involved reviewing patient records and assigning appropriate terms based on clinical manifestations, ensuring standardized phenotypic representation across the cohort. Results generated 164 patient-specific variant rankings per tool, consolidated into structured TSV tables containing genomic position (GRCh38), alleles, affected gene, and ranking positions.

2.3. Performance Evaluation

Results were compared with validated diagnoses. Performance was assessed using top-K recall metrics, measuring the percentage of patients for whom the validated causal variant appeared within the top K ranked variants for each method (K = 1, 5, 10, 15, 20). Results were stratified by disease group and subtype.

3. Results

3.1. Dataset overview

Basic demographic, clinical, and genetic characteristics of the analyzed dataset (N=164) are shown in Table 1. A summary on HPOs and causal variants is shown in Figure 1.

Table 1. Demographic characteristics (N=164)

Characteristic	Distribution
Sex	Male: 55%, Female: 45%
Age	Mean: 45.3 ± 16.2, Range: 2-83
Race	White: 50.6%, Pardo: 32.5%, Black: 14.4%, Asian: 1.9%

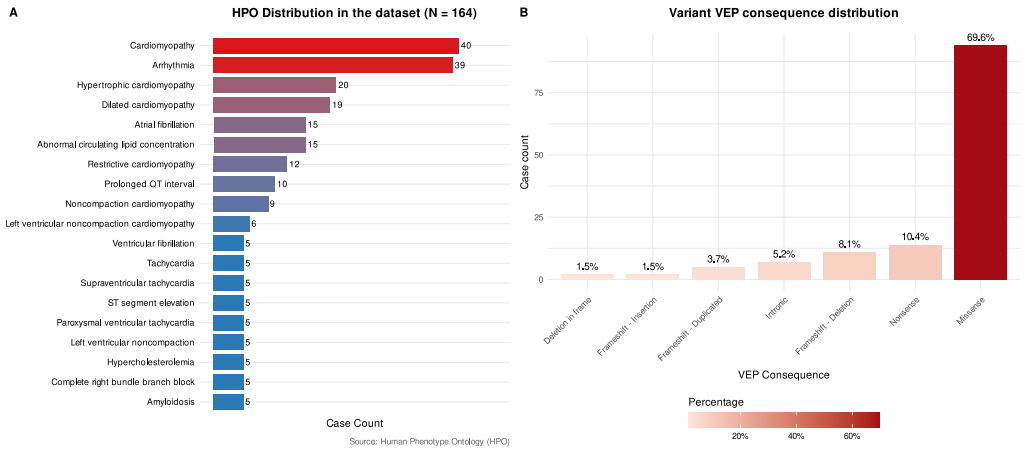

Figure 1. (A) HPO distribution; (B) Causal variant consequences

3.2. Benchmarking Results

We benchmarked all four prioritization tools on 164 whole-exomes with validated diagnoses. We obtained recall rate at top-K ranking for each method.

Figure 2 shows both Exomiser-based methods (hiPhive and PhenIX) outperformed the others, with PhenIX showing advantage from top-10. LIRICAL maintained 10-20% advantage over xRare.

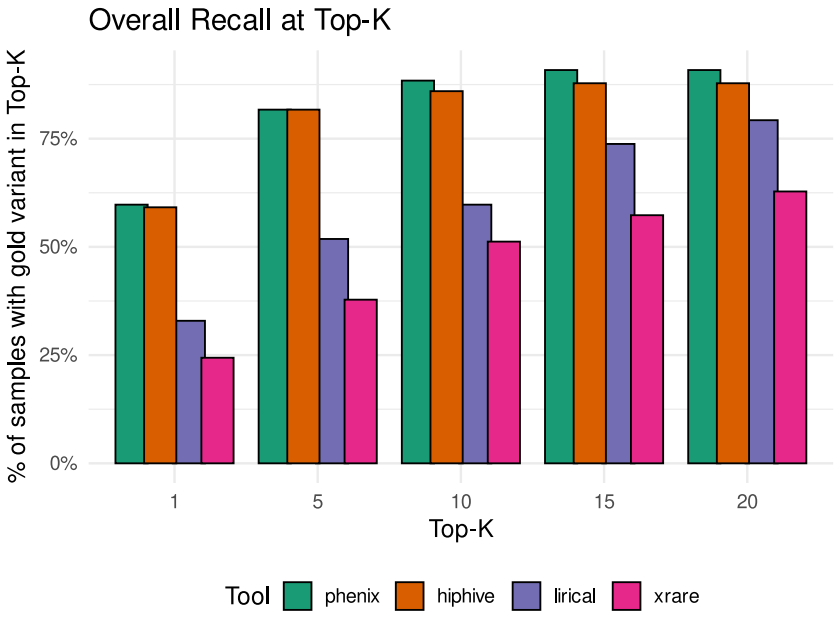

Figure 2. Top-K recall rates for all samples (N=164)

Figure 3 shows hiPhive, PhenIX, and xRare performed best for Dyslipidemias, with Exomiser methods achieving 100% recall at top-10. xRare showed superior accuracy for Dyslipidemias compared to Arrhythmias and Cardiomyopathies. LIRICAL performed better for Arrhythmias and Cardiomyopathies.

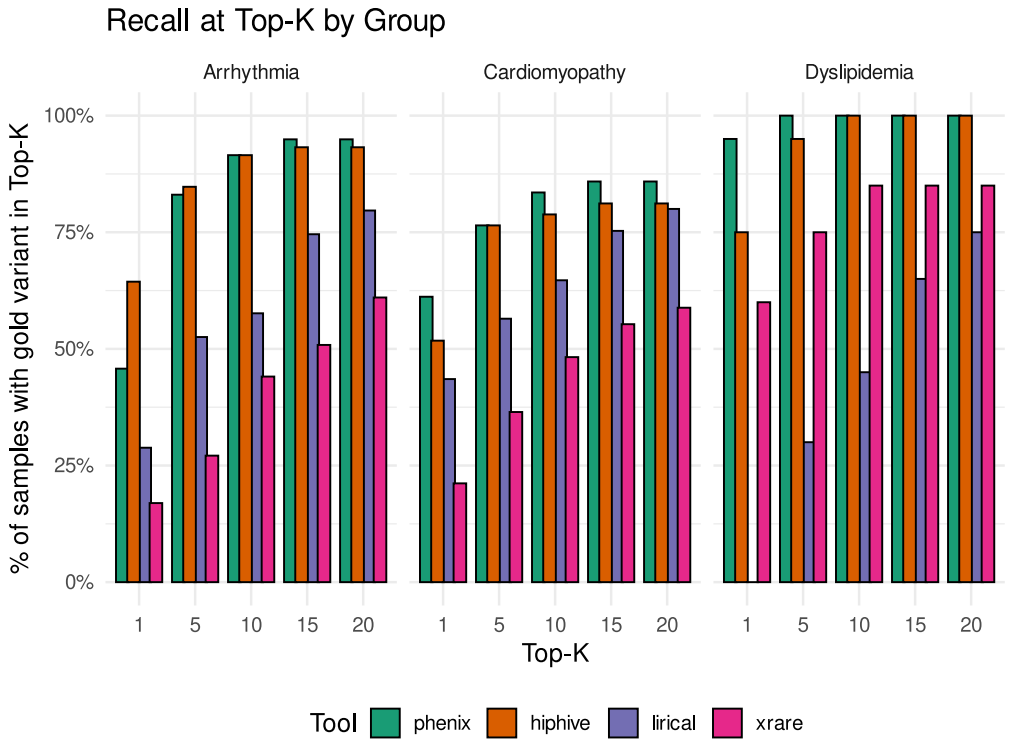

Figure 3. Top-K recall by diagnosis group (Arrhythmia N=59, Cardiomyopathy N=85, Dyslipidemia N=20)

Figure 4 shows Exomiser methods dominated overall, with PhenIX displaying uniform performance. hiPhive performed well in Hypertrophic Cardiomyopathy and Dyslipidemias. LIRICAL excelled in Amyloidosis (90% at top-1) and Restrictive Cardiomyopathies. xRare performed worst overall but surpassed LIRICAL in Dilated Cardiomyopathy and Familial Hypercholesterolemia.

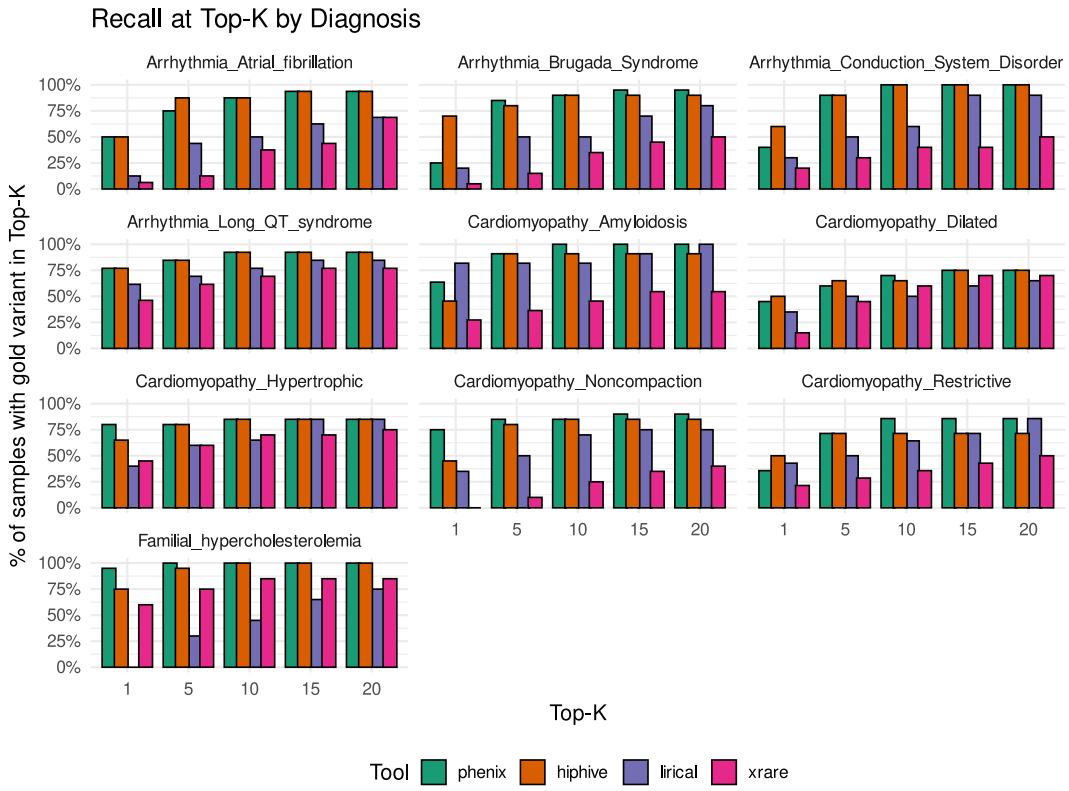

Figure 4. Top-K recall by diagnosis subgroup

4. Discussion

We systematically evaluated four variant prioritization tools - hiPhive, PhenIX, LIRICAL, and xRare - in a Brazilian cohort of 164 patients with genetic cardiac conditions, leveraging validated diagnoses as a gold standard. The results demonstrated that hiPhive and PhenIX generally outperformed LIRICAL and xRare, with PhenIX exhibiting superior recall rates at higher top-K rankings (e.g., top-10 to top-20). xRare struggled with disease groups like arrhythmias and cardiomyopathies.

When looking at disease subgroups, hiPhive and PhenIX were also superior over LIRICAL and xRare, with PhenIX showing the most uniform performance across all subgroups. hiPhive performed particularly well in cases such as Hypertrophic Cardiomyopathy and Dyslipidemias, while LIRICAL stood out in ranking variants from patients with Amyloidosis (nearly 90% recall at top-1) and Restrictive Cardiomyopathies, where it even surpassed hiPhive at top-20. Although xRare was the overall worst-performing method, it outperformed LIRICAL in a few specific contexts, such as Dilated Cardiomyopathy and Familial Hypercholesterolemia. Still, across all tools, Dilated Cardiomyopathy cases showed the lowest recall rates overall, suggesting a possible advantage to an integrative approach in this subgroup, for example.

Our findings underscore the importance of integrating multiple genomic prioritization tools to enhance diagnostic precision. Specifically, combining hiPhive and PhenIX under a single Machine Learning ensemble model may lead to superior performance compared to their isolated use, significantly aiding geneticists in achieving more accurate and

timely diagnoses. This integrative strategy is particularly relevant in the context of public healthcare systems such as Brazil's SUS (*Sistema Único de Saúde*), where access to specialized genetic centers may be limited. Patients in remote or underserved regions would benefit the most, ultimately improving clinical outcomes and reducing diagnostic delays.

5. Acknowledgments

This study was funded by CNPq project 444224/2023-8 (*Saúde de Precisão*, Chamada 16/2023), entitled *"Plataforma Web para priorização de variantes genéticas de doenças cardiovasculares, baseada em meta-aprendizado em população brasileira, integrando vias metabólicas e perfil clínico"*.

References

Köhler, S., Gargano, M., Matentzoglu, N., Carmody, L. C., Lewis-Smith, D., Vasilevsky, N. A., Danis, D., Balagura, G., Baynam, G., Brower, A. M., et al. (2021). The human phenotype ontology in 2021. *Nucleic acids research*, 49(D1):D1207–D1217.

Li, Q., Zhao, K., Bustamante, C. D., Ma, X., and Wong, W. H. (2019). Xrare: a machine learning method jointly modeling phenotypes and genetic evidence for rare disease diagnosis. *Genetics in Medicine*, 21(9):2126–2134.

Robinson, P. N., Köhler, S., Oellrich, A., Wang, K., Mungall, C. J., Lewis, S. E., Washington, N., Bauer, S., Seelow, D., Krawitz, P., et al. (2014). Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome research*, 24(2):340–348.

Robinson, P. N., Ravanmehr, V., Jacobsen, J. O., Danis, D., Zhang, X. A., Carmody, L. C., Gargano, M. A., Thaxton, C. L., Karlebach, G., Reese, J., et al. (2020). Interpretable clinical genomics with a likelihood ratio paradigm. *The American Journal of Human Genetics*, 107(3):403–417.

Smedley, D., Jacobsen, J. O., Jäger, M., Köhler, S., Holtgrewe, M., Schubach, M., Siragusa, E., Zemojtel, T., Buske, O. J., Washington, N. L., et al. (2015). Next-generation diagnostics and disease-gene discovery with the exomiser. *Nature protocols*, 10(12):2004–2015.