POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

A network medicine approach for drug repurposing in arrhythmogenic cardiomyopathy

Laurea Magistrale in Mathematical Engineering - Ingegneria Matematica

Author: Aurora Vido

Advisor: Prof. Paolo Zunino

Co-advisors: Dr. Cristina Banfi, Ph.D., Dr. Elena Sommariva Ph.D.

Academic year: 2024-2025

1. Introduction

Arrhythmogenic Cardiomyopathy (ACM) is an inherited heart muscle disease characterized by a progressive loss of cardiomyocytes and their replacement with fibrofatty tissue, predisposing to life-threatening ventricular arrhythmias and sudden cardiac death, especially in young patients and athletes [1]. The genetic architecture of ACM is centered on mutations in genes encoding desmosomal proteins, which form the mechanical junctions between cardiomyocytes. Despite the advances in genetic characterization and clinical trials, there is no curative therapy for ACM yet, and the current treatments are mostly symptomatic and preventive [1]. This underlines the need for novel approaches that can modify disease progression at a molecular level. Genetic heterogeneity suggests that ACM is not a simple monogenic disorder, rather, it is a disease of network failure: the various genetic insults converge on a common set of downstream cellular signaling pathways that make ACM an ideal candidate for a network medicine approach [1]. Traditional drug discovery focuses on the "one target, one drug" paradigm, which has proved to be highly inadequate for complex disorders like ACM [2, 3]. However, network medicine offers a paradigm shift by conceptualizing diseases as the outcome of network perturbations rather than single-gene defects, where pathological processes emerge when "disease modules" within the interactome are perturbed [2]. An additional powerful strategy is drug repurposing, referring to the use of approved drugs for treating diseases beyond their original medical indication [4]. If diseases are linked by overlapping molecular pathways, an existing drug can retain therapeutic efficacy across both [2]. Within the network medicine framework, this is applied by identifying drugs whose targets are located within, or in close proximity to, the disease module of interest [3]. The main objective of this thesis is to create and validate an integrated computational tool for drug repurposing in arrhythmogenic cardiomyopathy. Specifically, this work aims to integrate transcriptomic and protein-protein interaction data to apply a network-based algorithm and identify potential repurposable drug candidates that could be further validated *in vitro*.

2. Data description

This analysis employed the human protein-protein interaction (PPI) network from the

STRING database [5], which considers functional associations alongside physical protein binding. Each interaction is assigned a combined confidence score, and by applying a strict filtering criterion, only interactions with a combined score ≥ 0.7 were retained, prioritizing solid links supported experimentally or by computational methods.
A set of genes known to be causally associated with ACM was manually curated from clinical databases [6, 7]. This gene set, denoted as the disease module S, constitutes the "seed" nodes for the network propagation analysis, representing the core pathogenic perturbations.
Drug-target information was obtained from DrugBank [8]. Target UniProt IDs were converted into their corresponding Entrez gene IDs. Withdrawn, illicit, veterinary drugs and compounds that are toxic to humans were excluded. The final dataset consisted of 5,089 drugs.

3. Methods

3.1. Random Walk with Restart

The Random Walk with Restart (RWR) algorithm is a powerful network propagation method used to measure the functional proximity and influence between nodes in a network [9]. In the context of drug repurposing, the core intuition is that the influence of a disease diffuses through the network, and a drug is considered a promising candidate if its targets lie within the "hot zones", i.e., regions that are highly influenced by the disease genes [10, 11].
The method requires three inputs: the human interactome (specifically, the transition matrix W), the disease module S (seed nodes) and the restart probability γ. This hyperparameter was obtained through Leave-One-Out Cross-Validation (LOO-CV), which identified the value that best recovered the known disease genes when they were left out of the seed set. This optimization yielded an optimal damping factor of 0.60, and thus a restart probability of 0.40.
The iterative formula combines the random walk on the graph and the restart to the seed nodes:

$$\vec{p}_{t+1} = (1-\gamma)W\vec{p}_t + \gamma\vec{p}_0 \qquad (1)$$

where $\vec{p}_t$ is the probability vector at step t, W is the transition matrix and γ is the restart probability. This iteration continues until the vector converges to a steady state $\vec{p}_\infty$, i.e., when the difference between $\vec{p}_{t+1}$ and $\vec{p}_t$ (measured by the L1-norm) falls below a small threshold (e.g., 10^{-6}). The vector $\vec{p}_\infty$ represents a comprehensive ranking of all proteins based on their relevance with respect to the ACM disease module.

3.2. Permutation test

The interactome was modeled as a weighted undirected graph, where the weight w_{ij} of the edge connecting gene i and gene j corresponds to the STRING combined confidence score. The RWR algorithm was then executed on the Largest Connected Component (LCC) of the network, which was extracted to ensure that all proteins were reachable from all other nodes. The RWR algorithm yielded a steady-state probability vector $\vec{p}_\infty$ which assigned a disease relevance score to each gene. The raw scores were then log-transformed to mitigate extreme skewness caused by the power-law distribution typical of scale-free networks. In order to rigorously quantify the significance of each drug, a degree-preserving permutation test was employed. For each drug D with a set of targets T_D, the observed "proximity" score was computed as the mean log-transformed RWR score of its targets:

$$S_{obs}(D) = \frac{1}{|T_D|}\sum_{t\in T_D} logScore_t \qquad (2)$$

To construct a null distribution, 10,000 sets of random targets for each drug were generated. It is important to note that these random sets were not extracted uniformly from the network; instead, each real target was replaced by a random node with equal degree. The nodes were grouped according to their degree and random samples were drawn from the corresponding groups. This ensures that statistical significance reflects specific proximity to the disease rather than generic network centrality.
For each drug, a Z-score was computed:

$$Z_D = \frac{S_{obs} - \mu_{rand}}{\sigma_{rand}} \qquad (3)$$

where S_{obs} is the observed proximity score of a drug, μ_{rand} and σ_{rand} are the mean and standard

deviation of the scores from the 10,000 permutations. An empirical p-value was calculated as the fraction of permutations achieving a score higher than the observed one. The p-values were subsequently adjusted using the Benjamini-Hochberg (FDR) procedure.

3.3. Network separation analysis

To better understand the mechanism of action of each drug, the network separation metric s_{AB} [12] was computed. It measures the topological relationship between the drug target module and the disease module, in particular it quantifies how distinct the two groups of nodes are:

$$s_{AB} = \langle d_{AB} \rangle - \frac{\langle d_{AA} \rangle + \langle d_{BB} \rangle}{2} \tag{4}$$

where $\langle d_{AB} \rangle$ is the average shortest distance between the drug targets and the disease genes, $\langle d_{AA} \rangle$ and $\langle d_{BB} \rangle$ are the average internal distances within the drug target set and the disease set respectively. The sign and magnitude of s_{AB} distinguish between three topological classes:

- $s_{AB} < 0$ **(Overlap):** the drug targets and disease genes form overlapping neighborhoods.
- $s_{AB} \approx 0$ **(Proximity):** drugs act in the immediate proximity of the disease set.
- $s_{AB} > 0$ **(Separation):** the drug targets are topologically separated from the disease module.

3.4. Phenotype context scoring

A Context Score was defined to refine the prioritization, exploiting the "guilt by association" and "shared gene" hypotheses, which assume that phenotypic or genetic similarities between diseases indicate shared pathobiological mechanisms [2, 3]. Consequently, diseases that interact within the same network neighborhood or share genetic origins are likely to respond to similar pharmacological treatments [2].
The Jaccard Similarity Index (J) [12, 13] was computed to rank more than 1000 diseases based on genetic similarity to ACM:

$$J(ACM, D_x) = \frac{|G_{ACM} \cap G_{D_x}|}{|G_{ACM} \cup G_{D_x}|} \tag{5}$$

where G_{ACM} is the set of ACM seed genes and G_{D_x} is the gene set for disease x. Diseases with $J > 0$ (sharing at least one gene) were selected as the phenotypic context, yielding a total of 844 diseases.
This metric enabled to rank the entire Phenopedia database based on genetic similarity to ACM. The top 10 diseases identified are listed in Figure 1. Notably, the ranking placed *Bundle-Branch Block* and *Brugada Syndrome* at the top, reflecting the strong electrical component of ACM. Furthermore, the inclusion of *Dilated Cardiomyopathy* and *Sudden Cardiac Death* confirms that the Jaccard Index correctly captures the structural degeneration and arrhythmic risk of the disease.

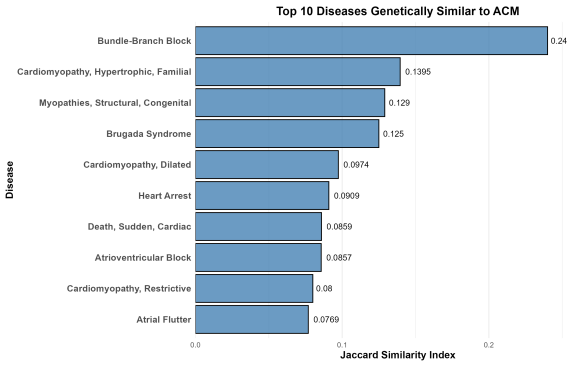

Figure 1: Top 10 diseases genetically similar to ACM.

Consequently, a weight w_g was assigned to each gene in the interactome:

$$w_g = \frac{\sum_{D_x \in \text{Context}} (\mathbb{1}(g \in D_x) \cdot J(ACM, D_x))}{w_{max}} \tag{6}$$

where $\mathbb{1}(g \in D_x)$ indicates the presence of the gene in disease x, and w_{max} is a normalization factor corresponding to the maximum observed score.
Lastly, the Context Score was computed for each drug D as the average weight of its targets, rewarding high specificity for the cardiac context:

$$S_{context}(D) = \frac{1}{|T_D|} \sum_{t \in T_D} w_t \tag{7}$$

where T_D is the set of targets for drug D, and $|T_D|$ is the total number of targets.

4. Results

The prioritization strategy integrated the statistical robustness of the degree-preserving permutation test with the biological specificity of

the phenotype-driven context score, and a dual filtering criterion was applied to ensure the selection of high-confidence candidates:

1. Statistical significance: an adjusted p-value < 0.05 from the permutation test.
2. Phenotypic consistency: a Context Score > 0. While this threshold technically acts as a strict consistency filter to exclude drugs with no relevance to the cardiac phenotypic context, a score of exactly zero is highly uncommon for candidates that have already demonstrated topological significance. Therefore, the primary biological utility of this metric is not mere exclusion, but rather establishing the final hierarchical ranking of the candidates based on their specific phenotypic context.

This procedure identified 118 promising candidates. The top 20 are presented in Table 1, ranked by Context Score.

Drug Name	Z-Score	Adj. P-val	Context Score	s_{AB}
Lisinopril	4.80	1.82×10^{-2}	0.58	1.10
Enalaprilat	4.81	2.24×10^{-2}	0.52	1.29
Ramipril	4.77	2.88×10^{-2}	0.52	1.29
Candoxatril	4.54	4.39×10^{-2}	0.52	1.12
Ilepatril	4.51	4.02×10^{-2}	0.52	1.12
Omapatrilat	4.51	1.82×10^{-2}	0.52	1.12
Gallopamil	5.67	7.83×10^{-3}	0.38	0.60
Hydrochlorothiazide	4.01	2.88×10^{-2}	0.35	0.74
Captopril	4.33	1.32×10^{-2}	0.31	0.82
Propafenone	4.49	2.24×10^{-2}	0.26	0.47
Dyclonine	4.73	2.88×10^{-2}	0.20	0.74
Hexylcaine	4.81	3.21×10^{-2}	0.20	0.74
Azimilide	4.29	3.21×10^{-2}	0.18	0.87
Procainamide	4.35	2.88×10^{-2}	0.18	0.11
Aprindine	5.30	7.83×10^{-3}	0.18	0.31
Flecainide	7.04	7.83×10^{-3}	0.17	0.47
Vernakalant	6.02	7.83×10^{-3}	0.14	0.77
Neratinib	3.42	4.74×10^{-2}	0.13	0.59
Canertinib	3.80	2.65×10^{-2}	0.12	0.62
Carvedilol	4.78	7.83×10^{-3}	0.12	0.58

Table 1: Top significant drug repurposing candidates ranked by Context Score.

Furthermore, the relationship between statistical significance (Z-score), Context Score and topological location (Network Separation s_{AB}) is illustrated in Figure 2. Interestingly, the analysis revealed that no approved drugs directly overlap with the ACM disease module ($s_{AB} < 0$); instead, the majority of the significant candidates cluster in the "Proximal" region ($s_{AB} \approx 0.2-1$), and so they target proteins that are neighbors of the disease genes.

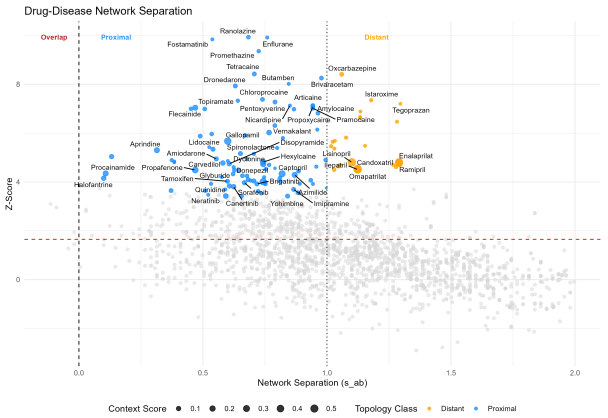

Figure 2: Network-based drug repurposing map.

The ranking successfully recovered widely used antiarrhythmic drugs (e.g., *Flecainide*, *Propafenone*) in the proximal zone. The distant zone revealed a cluster of ACE inhibitors (e.g., *Lisinopril*, *Ramipril*), which reached the top positions thanks to the RWR algorithm's ability to identify genes relevant to the disease even over greater distances. Moreover, to assess whether the ranking methodology is meaningful, a validation analysis was conducted using a "gold standard" set of drugs manually curated (e.g., β-blockers, antiarrhythmics) [14–16]. The performance was quantified using Receiver Operating Characteristic (ROC) curve analysis. The model achieved an Area Under the Curve (AUC) of 0.861, suggesting that this network-based ranking performs significantly better than random chance and successfully enriches for known cardiac relevant drugs.

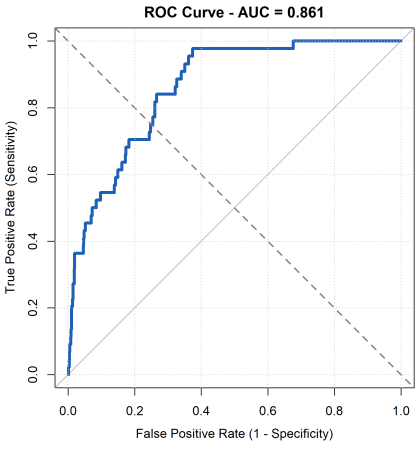

Figure 3: ROC curve validating the performance of the drug ranking methodology.

5. Topological analysis

In order to better understand the properties of the identified repurposing candidates, a topological analysis was performed on the target genes of the 118 statistically significant drugs. Two main questions were addressed:

1. Are the targets of significant drugs topologically distinct from the targets of non-significant drugs?
2. Which topological features best predict the efficacy of a drug target?

The targets of significant drugs ($pvalue_{adj} < 0.05$) were compared against a background set of non-significant drugs ($pvalue_{adj} > 0.8$). Five key metrics were analyzed: degree, betweenness centrality, eigenvector centrality, clustering coefficient and minimum distance to disease.
First, a univariate Wilcoxon rank-sum test (see Table 2) showed that the most significant difference lies in the minimum distance to disease, where significant targets are closer to the disease module ($p < 10^{-16}$). Significant targets also exhibit higher betweenness and eigenvector centrality, but a lower clustering coefficient. Interestingly, the degree showed no significant difference ($p = 0.315$).

Metric	Median (Sig)	Median (Non-Sig)	P-value	Signif
Degree	30.0	31.0	3.15×10^{-1}	ns
Betweenness	1.51×10^{-4}	1.12×10^{-4}	4.46×10^{-4}	***
Eigenvector	5.60×10^{-4}	4.38×10^{-4}	2.03×10^{-3}	**
Clustering Coeff.	0.27	0.32	4.66×10^{-8}	***
Dist. to Disease	3.09	3.60	1.98×10^{-41}	***

Table 2: Comparison between targets of significant and non-significant drugs. Significance levels: ** $p < 0.01$, *** $p < 0.001$, ns = not significant.

However, topological features are often highly correlated, for instance, nodes with a large number of connections (high degree) naturally tend to have high betweenness, potentially introducing multicollinearity into statistical models. Indeed, the Spearman correlation matrix in Figure 4 clearly shows these strong correlations between the metrics. In order to separate the independent contribution of each metric, a multivariate logistic regression model was implemented to see which features best predict the efficacy of a drug target. The Variance Inflation Factor (VIF) was computed for all predictors to ensure the stability of the regression coefficients. All the features yielded VIF values lower than 5.0, well below the critical threshold, confirming that, while correlations exist, they do not reach the level of severe multicollinearity that would compromise the model. Consequently, the resulting Odds Ratios (see Table 3) can be interpreted as the unique predictive power of each topological feature, independent of the others.

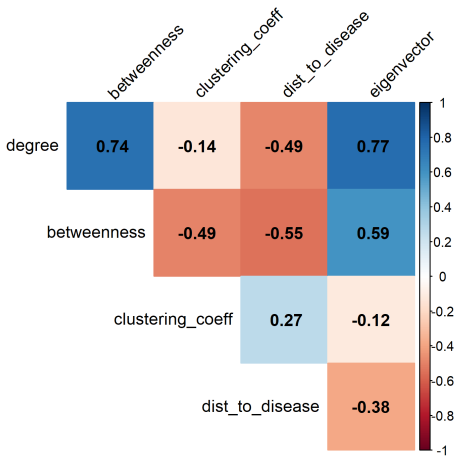

Figure 4: Spearman correlation matrix of network metrics.

The multivariate logistic regression predicts the probability of a target being "significant" based on the scaled topological features. The model is defined as:

$$\begin{aligned}\text{logit}(p) =& \beta_0 + \beta_1 \cdot \text{deg} + \beta_2 \cdot \text{betw} \\ &+ \beta_3 \cdot \text{eigen} + \beta_4 \cdot \text{clust} + \beta_5 \cdot \text{dist}\end{aligned} \tag{8}$$

The results of the multivariate analysis (Table 3) clarify these relationships, isolating the independent effects masked in the univariate tests:

- **Distance to disease ($OR = 0.448$):** this remains the strongest predictor. Proximity to the disease module is the primary driver of repurposing potential.
- **The centrality paradox:** the degree is a negative predictor ($OR = 0.589$) when controlling for other centralities. This confirms that the degree-preserving permutation strategy successfully penalized generic hub-genes. Conversely, betweenness ($OR =$

1.454) and eigenvector centralities ($OR = 1.202$) are significant positive predictors.
- **Clustering coefficient:** once controlled for centrality and distance, local clustering becomes non-significant ($p = 0.769$).

Predictor	Odds Ratio	2.5 %	97.5 %	Signif
(Intercept)	0.622	0.569	0.680	***
Degree (Scaled)	0.589	0.520	0.667	***
Betweenness (Scaled)	1.454	1.209	1.750	***
Dist. to Disease (Scaled)	0.448	0.399	0.503	***
Eigenvector (Scaled)	1.202	1.098	1.314	***
Clustering (Scaled)	0.986	0.895	1.086	ns

Table 3: Logistic regression results (Odds Ratios). An $OR > 1$ indicates a positive association with target significance, while $OR < 1$ indicates a negative association. Significance levels: ** $p < 0.01$, *** $p < 0.001$, ns = not significant.

These results suggest that the ideal target is not a "super-hub" with the highest number of raw connections; instead, the analysis favors nodes with "strategic centrality", i.e. genes that act as bridges controlling information flow or genes connected to other influential nodes, provided they are located in the vicinity of the disease module.

6. Community analysis

Since diseases are often caused by the dysfunction of specific biological modules rather than scattered genes, the following analysis focused on the disease module to investigate its internal heterogeneity.

6.1. Expanded disease module

The concept of "disease module" was expanded beyond the initial seed genes to include all genes with high Random Walk with Restart (RWR) scores. The boundaries of this new module were established using the "elbow method", identifying a cutoff at the point of maximum curvature. This procedure defined a final ACM Disease Module consisting of 65 nodes (Figure 5).

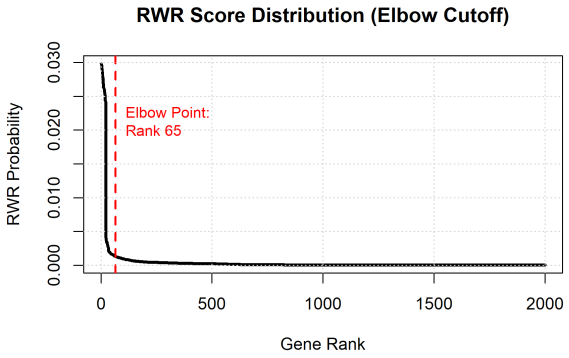

Figure 5: Ranked RWR scores for the definition of the ACM Disease Module. Cutoff at rank 65 (red line).

Crucially, a cross-reference analysis reveals that 9 of the 118 significant drug candidates, including *Dronedarone*, *Ramipril*, and *Gallopamil*, act directly upon these high-score genes.

6.2. Functional partitioning

In order to better understand the biological heterogeneity of the disease, the Louvain algorithm [17, 18] was applied to this specific subgraph of 65 nodes, partitioning it into 4 distinct communities, which were subjected to functional enrichment analysis using Metascape [19].
Module 1 is reported as a representative case study (Figure 6).

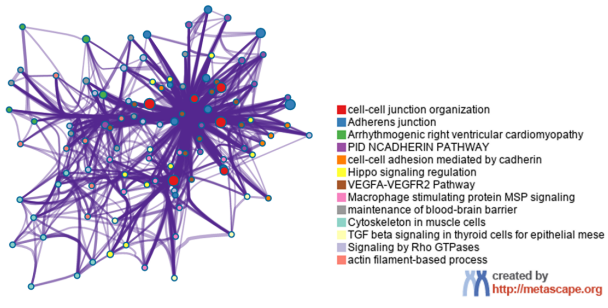

Figure 6: Network of enriched terms: colored by cluster ID, where nodes that share the same cluster ID are typically close to each other [19].

This visualization provides an overview of the mechanisms of the disease. Each node in the graph represents an enriched biological term (e.g., a pathway) and the edges connect terms that share a significant number of genes. The clustering of these terms (indicated by color) shows that this module alters coherent biologi-

cal pathways, in particular, the dense clustering of terms related to "cell-cell junction organization" and "adherent junctions" (red/blue clusters) confirms that the primary defect in ACM lies in structural integrity. However, the connection of these structural nodes to signaling pathways such as "Hippo" and "TGF-β" (yellow clusters) provides the functional link that explains how mechanical detachment translates into the fibro-adipose remodeling observed in patients.

7. Transcriptomic validation

A transcriptomic validation analysis was performed using the Connectivity Map (CMap) [20] to complete the network-based prioritization strategy. This orthogonal approach aims to identify drugs capable of reversing the global gene expression signature associated with ACM, a concept known as "signature reversion".
The Connectivity Map (CMap) is a large-scale catalog of cellular responses to chemical, genetic and disease perturbations [20]. The core principle of CMap is "pattern matching": by comparing the gene expression signature of a disease, i.e., a list of up- and down-regulated genes, against the signatures induced by thousands of small molecules, it is possible to identify compounds that induce an opposing transcriptional profile. This is quantified by a "Connectivity Score" (or τ-score), ranging from -100 to +100. A score approaching -100 indicates a strong potential to reverse the disease phenotype [20]. In this work, drugs with a score lower than -80 were selected as plausible candidates for reversing the disease signature.

7.1. ACM Disease Signatures

Five distinct transcriptomic datasets were selected from the literature to ensure a robust validation that captures the heterogeneity of ACM. These datasets represent different stages of the disease, different biological materials and different detection technologies.

- **Dataset 1 (Rouhi et al. 2022 [21]):** represents the early, arrhythmic phase (Right Ventricular biopsies).
- **Dataset 2 (Rainer et al. 2018 [22]):** characterized the transcriptome of Cardiac Stromal Cells (CStCs), highlighting the fibrotic drive.
- **Dataset 3 (Gaertner et al. 2012 [23]):** reflects the terminal stage of the disease (explanted hearts).
- **Dataset 4 (De Bortoli et al., 2023 [24]):** analyzes hiPSC-CMs, distinguishing between pathogenic manifestation and simple mutation carriage.
- **Dataset 5 (Lippi et al., 2023 [25]):** examines Cardiac Mesenchymal Stromal Cells (CMSCs), highlighting mitochondrial homeostasis and metabolic alterations.

7.2. Consensus scoring system

A "consensus scoring system" was developed to integrate the network-based findings with this transcriptomic evidence. The 118 significant drug candidates identified by the RWR algorithm were first cross-referenced against the CMap output lists (Table 4).

Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
Carbamazepine	Gallopamil	Dronedarone	Carvedilol	Spironolactone
Isradipine	Disopyramide	Tetracaine	Disopyramide	Ponatinib
Nimodipine	Lacidipine	Carbamazepine	Isradipine	Amiloride
Topiramate	Dalfampridine	Ritodrine	Verapamil	Tivozanib
Fostamatinib	Digoxin	Isradipine	Topiramate	
Sorafenib	Canertinib	Nifedipine	Lacidipine	
Canertinib	Flunarizine	Zonisamide	Istaroxime	
Imipramine		Trimebutine	Nicardipine	
		Nitrendipine	Fostamatinib	
		Digoxin	Amiodarone	
		Regorafenib	Amitriptyline	
		Phenytoin	Tivozanib	
		Ramipril		
		Nintedanib		
		Lamotrigine		
		Imipramine		
		Flunarizine		

Table 4: Intersection of RWR prioritized drugs with CMap revertant signatures.

A discrete score was then assigned to each candidate based on the number of transcriptomic signatures it successfully reversed:

$$S_{consensus}(D) = \sum_{i=1}^{5} \mathbb{I}(D \in \text{CMap}_i) \qquad (9)$$

where CMap_i is the set of drugs with a significant negative connectivity score for Dataset i and $\mathbb{I}$ is the indicator function (1 if true, 0 if false). This multi-layer filtering ensures that candidates are not only located in the network neighborhood but are also functionally capable of modulating global gene expression.

The analysis highlighted *Isradipine* ($S = 3$) as the most robust novel candidate, as it is able to reverse three disease signatures. The screening successfully identified treatments currently used or under investigation for ACM (e.g., *Amiodarone*, *Carvedilol*, *Ramipril*, *Spironolactone*) in the subset with $S = 1$, serving as a strong biological validation of the network model. Among novel candidates, *Fostamatinib* ($S = 2$) offers a targeted approach to structural remodeling, particularly inflammation and fibro-adipose replacement [26, 27].

Drug Candidate	Score	D1	D2	D3	D4	D5
High Confidence (Score ≥ 2)						
Isradipine	**3**	✓		✓	✓	
Canertinib	2	✓	✓			
Carbamazepine	2	✓		✓		
Digoxin	2		✓	✓		
Disopyramide	2		✓		✓	
Flunarizine	2		✓	✓		
Fostamatinib	2	✓			✓	
Imipramine	2	✓		✓		
Lacidipine	2		✓		✓	
Tivozanib	2				✓	✓
Topiramate	2	✓			✓	
Validation Hits (Score 1)						
Amiodarone	1				✓	
Carvedilol	1				✓	
Ramipril	1			✓		
Spironolactone	1					✓
Verapamil	1				✓	

Table 5: Candidates with Score ≥ 2 and selected clinically relevant candidates with Score 1. Checkmarks indicate successful signature reversion.

8. Conclusions

This study developed and validated a robust computational framework for drug repurposing based on the principles of network medicine, confirming the central hypothesis that the effects of ACM genetic defects propagate through specific modules of the human interactome.
The proposed prioritization pipeline based on network propagation, degree-preserving permutation test and phenotype-based context scoring identified 118 statistically significant drug candidates. The accuracy of this ranking was validated with a ROC-AUC analysis, which achieved an AUC of 0.861 and successfully recovered drugs such as antiarrhythmics (*Amiodarone*), beta-blockers (*Carvedilol*) and ACE inhibitors (*Ramipril*). This approach was further strengthened by integrating CMap/L1000 data to identify drugs capable of reversing the pathological gene expression signatures across multiple patient datasets. Notably, topological analysis revealed that therapeutic efficacy is driven by targeting "network bridges" characterized by high betweenness and eigenvector centrality, rather than generic high-degree hubs. These findings provide a direct foundation for experimental studies at Centro Cardiologico Monzino (IRCCS), where the top candidates will be tested *in vitro* to assess their efficacy to restore the cellular phenotype.
Beyond the necessary biological validation of the main candidates, the pipeline has also the potential to evolve into a patient-specific diagnostic tool. By mapping the specific mutations and transcriptomic profile of an individual patient onto the network, this framework could directly predict the most effective targeted therapy for that individual, thus moving from a generalized therapeutic approach to precision medicine.
Furthermore, the pipeline represents an adaptable and reproducible methodology, documented and available on GitHub `https://github.com/AuroraVido/NetMedWalkeR`. This open-source repository allows the scientific community to use the method for different input data or even improve the algorithm to better tackle the challenges of modern drug discovery.

References

[1] Judge DP Corrado D, Basso C. Arrhythmogenic cardiomyopathy. *Circulation research*, pages 784–802, 2017.

[2] Loscalzo J Barabási A-L, Gulbahce N. Network medicine: a network-based approach to human disease. *Nat Rev Genet*, 2011.

[3] Loscalzo J. Chan, S. Y. The emerging paradigm of network medicine in the study of human disease. *Circulation research*, 2012.

[4] Park K. A review of computational drug re-

purposing. *Translational and clinical pharmacology*, 2019.

[5] Kirsch R. Koutrouli M. Nastou K. Mehryary F. Hachilif R. Gable A. L. Fang T. Doncheva N. T. Pyysalo S. Bork P. Jensen L. J. von Mering C. Szklarczyk, D. The string database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, 51(D1):D638–D646, 2023.

[6] Muin J. Khoury Wei Yu, Melinda Clyne and Marta Gwinn. Phenopedia and genopedia: Disease-centered and gene-centered views of the evolving knowledge of human genetic associations. *Bioinformatics*, 2010.

[7] Chitipiralla S. Brown G. R. Chen C. Gu B. Hart J. Hoffman D. Jang W. Kaur K. Liu C. Lyoshin V. Maddipatla Z. Maiti R. Mitchell J. O'Leary N. Riley G. R. Shi W. Zhou G. Schneider V. Maglott D. ... Kattman B. L. Landrum, M. J. Clinvar: improvements to accessing data. *Nucleic acids research*, 2020.

[8] Klinger CM et al. Knox C, Wilson M. Drugbank 6.0: the drugbank knowledgebase for 2024. *Nucleic Acids Res*, 2024.

[9] Duc-Hau Le. Random walk with restart: A powerful network propagation algorithm in bioinformatics field. pages 242–247, 2017.

[10] Horn D Robinson PN Köhler S, Bauer S. Walking the interactome for prioritization of candidate disease genes. *Am J Hum Genet*, 2008.

[11] Alberto Valdeolivas, Laurent Tichit, Claire Navarro, Sophie Perrin, Gaëlle Odelin, Nicolas Levy, Pierre Cau, Elisabeth Remy, and Anaïs Baudot. Random walk with restart on multiplex and heterogeneous biological networks. *Bioinformatics*, 35(3):497–505, 07 2018.

[12] Sharma A. Kitsak M. Ghiassian S. D. Vidal M. Loscalzo J. Barabási A. L. Menche, J. Uncovering disease-disease relationships through the incomplete interactome. *Science*, 2015.

[13] Valle D et al. Goh K-I, Cusick ME. The human disease network. *Proc Natl Acad Sci*, 2007.

[14] Calkins H. Gaine, S. P. Antiarrhythmic drug therapy in arrhythmogenic right ventricular cardiomyopathy. *Biomedicines*, 11(4), 2023.

[15] Gerstenfeld E. P. Svetlichnaya Y. Scheinman M. M. Ermakov, S. Use of flecainide in combination antiarrhythmic therapy in patients with arrhythmogenic right ventricular cardiomyopathy. *Heart rhythm*, 14(4):564–569, 2017.

[16] Wu L. Zheng L. Liu S. Sheng L. Liu L. Zhu Z. Yao Y. Tu, B. Angiotensin-converting enzyme inhibitors/angiotensin receptor blockers: Anti-arrhythmic drug for arrhythmogenic right ventricular cardiomyopathy. *Frontiers in cardiovascular medicine*, 2021.

[17] Vincent Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics Theory and Experiment*, 2008.

[18] G. Rajendran D. Dhanalakshmi. Comparative study of louvain, leiden, and infomap algorithms for community detection in complex networks. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 2025.

[19] Zhou B. Pache L. Chang M. Khodabakhshi A. H. Tanaseichuk O. Benner C. Chanda S. K. Zhou, Y. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications*, 2019.

[20] Narayan R. Corsello S. M. Peck D. D. Natoli T. E. Lu X. Gould J. Davis J. F. Tubelli A. A. Asiedu J. K. Lahr D. L. Hirschman J. E. Liu Z. Donahue M. Julian B. Khan M. Wadden D. Smith I. C. Lam D. Liberzon A. ... Golub T. R. Subramanian, A. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell*, 2017.

[21] Fan S. Cheedipudi S. M. Braza-Boïls A. Molina M. S. Yao Y. Robertson M. J. Coarfa C. Gimeno J. R. Molina P. Gurha P. Zorio E. Marian A. J. Rouhi, L. The ep300/tp53 pathway, a suppressor of the hippo and canonical wnt pathways, is activated in human hearts with arrhythmogenic cardiomyopathy in the absence of overt heart failure. *Cardiovascular research*, 2022.

[22] Meraviglia V. Blankenburg H. Piubelli C.-Pramstaller P. P. Paolin A. Cogliati E. Pompilio G. Sommariva E. Domingues F. S. Rossini A. Rainer, J. The arrhythmogenic cardiomyopathy-specific coding and non-coding transcriptome in human cardiac stromal cells. *BMC genomics*, 2018.

[23] Schwientek P. Ellinghaus P. Summer H.-Golz S. Kassner A. Schulz U. Gummert J. Milting H. Gaertner, A. Myocardial transcriptome analysis of human arrhythmogenic right ventricular cardiomyopathy. *Physiological genomics*, 2012.

[24] Meraviglia V. Mackova K. Frommelt-L. S. König E. Rainer J. Volani C. Benzoni P. Schlittler M. Cattelan G. Motta B. M. Volpato C. Rauhe W. Barbuti A. Zacchigna S. Pramstaller P. P. Rossini A. De Bortoli, M. Modeling incomplete penetrance in arrhythmogenic cardiomyopathy by human induced pluripotent stem cell derived cardiomyocytes. *Computational and structural biotechnology journal*, 2023.

[25] Maione A. S. Chiesa M. Perrucci-G. L. Iengo L. Sattin T. Cencioni C. Savoia M. Zeiher A. M. Tundo F. Tondo C. Pompilio G. Sommariva E. Lippi, M. Omics analyses of stromal cells from acm patients reveal alterations in chromatin organization and mitochondrial homeostasis. *International journal of molecular sciences*, 2023.

[26] Wu Y. Zhou C. Xie J.-Zhang Y. Yang X. Xiao J. Wang D. W. Shan C. Zhou X. Xiang Y. Yang B. Zhang, B. Hyperactivation of atf4/tgf-1 signaling contributes to the progressive cardiac fibrosis in arrhythmogenic cardiomyopathy caused by dsg2 variant. *BMC medicine*, 2024.

[27] Can G. Ayvaz S. Karaca-T. Pamuk G. E. Demirtas S. Tsokos G. C. Pamuk, O. N. Spleen tyrosine kinase (syk) inhibitor fostamatinib limits tissue damage and fibrosis in a bleomycin-induced scleroderma mouse model. *Clinical and experimental rheumatology*, 2015.