Doctoral Thesis

Head Office: Università degli Studi di Padova

Department of Cardiac, Thoracic, Vascular Sciences and Public Health

---

Ph.D. COURSE IN: Translational Specialistic Medicine “G.B. Morgagni”

CURRICULUM: Cardiovascular Sciences

SERIES: 38°

THESIS TITLE:

Machine Learning models to create an expert system to predict major cardiac adverse events in monitoring heart transplant patients

Thesis written with the financial contribution of Name of the of the European Union - Next Generation EU and I4 Consulting S.r.l.

Coordinator: Prof. Dario Gregori

Supervisor: Prof. Chiara Castellani

Co-Supervisor: Dr. Pietro Fanton

Ph.D. student: Diego Perazzolo

Table of contents

Doctoral Program Statement

The classification of this doctoral program falls under the PNRR 352 (code: 38-033-19-DOTT1433495-2235) scholarship framework, which requires the development of two integrated research projects. The main project of the PhD focuses on the development of machine learning and deep learning models aimed at classifying, predicting, or detecting major cardiac events in the clinical monitoring of heart transplant patients. This research is conducted within the academic environment of the University of Padua and is aligned with the goals of personalized medicine and clinical decision support within the national healthcare system (SSN). As part of the program, six months of research were conducted abroad at London South Bank University (LSBU) under the supervision of Professor Enrico Grisan, whose expertise in biomedical signal and image processing contributed to the refinement and advancement of my computational modeling and machine learning knowledges. In addition, the PNRR 352 scholarship mandates a period of 18 months dedicated to an industrial research project. In this case, the project has been carried out in collaboration with I4 Consulting S.R.L., a multidisciplinary engineering firm based in Padua, Italy, specializes in hydrological and hydraulic modeling. I have been actively involved in the development of artificial intelligence models and ad hoc software tools for hydrological engineering applications. Although the fields of application differ, the underlying methodological approach based on AI and data-driven modeling demonstrates a shared potential for cross-disciplinary impact and future integration into public service infrastructures, including healthcare.

Summary

Heart transplantation represents the most effective treatment for patients with end-stage heart failure. Despite advances in surgical procedures and immunosuppressive therapies, long-term graft survival remains threatened by major complications such as acute cellular rejection (ACR) and viral infections, especially Cytomegalovirus (CMV). Endomyocardial biopsy (EMB) is currently the clinical standard for monitoring graft rejection, but complementary molecular and computational approaches are emerging to support and refine diagnostic assessment. Recent advances in biomarker discovery are paving the way for non-invasive, accurate, and early detection of adverse conditions, combined with routine EMB. These developments are increasingly supported by computational approaches that enable the integration of molecular signatures into predictive models for improved graft surveillance. In this thesis, I developed and validated a series of computational models based on machine learning (ML) and deep learning (DL) techniques to improve diagnostic precision and predictive capability in the post-transplant follow-up, with a strong emphasis on interpretability including data-scarce scenarios. The first part of this thesis presents a systematic bootstrap analysis conducted on a large publicly available microarray dataset (E-MTAB-8026), aimed at evaluating a machine learning framework specifically designed for data-scarce scenarios with a focus on interpretability. The study investigates the impact of data augmentation and feature selection techniques on classification performance and model transparency in omics data. The proposed pipeline integrates synthetic data generation, L1-regularized logistic regression (LASSO), and kernel-based classifiers to jointly optimize predictive accuracy and features selection. This framework was used to simulate clinically relevant low-sample conditions. Building upon this foundation, the second part of the thesis introduces AugPred, a machine learning pipeline specifically designed for post-transplant patient follow-up. AugPred was applied to the classification of Acute Cellular Rejection versus Cytomegalovirus infection using miRNA expression profiles obtained from endomyocardial biopsies of transplanted patients. Despite the limited sample size (n=11), the pipeline demonstrates that, through targeted data augmentation and robust feature selection, it is possible to train high-performing classifiers even in highly data-scarce scenarios. Comparative analyses confirmed its superior performance and interpretability over baseline approaches, while pathway enrichment analysis validated the involvement of the selected miRNAs in immune and infection-related biological processes. In the third part of the thesis, a multilayer network approach was employed to model miRNA-mRNA interactions derived from endomyocardial biopsy (EMB) samples of control patients, as well as those with ACR and CMV infection. By applying an ensemble PageRank centrality strategy, we identified miRNAs with phenotype-specific regulatory roles, that act as modulator between different conditions. These miRNAs were further validated through pathway enrichment analysis (PEA) for investigating the biological meaning. This network-based framework highlighted the potential of centrality-driven models to uncover molecular signatures associated with rejection and infection processes. The fourth section is dedicated to time series modeling, initially developed in collaboration with an industrial partner in the hydrological domain. Through the comparison of ARIMAX, LSTM recurrent neural network, and physical based models, I explored techniques for modeling sequential data, forecasting strategies, and the influence from external variables. The methodological insights gained were later transferred to the biomedical domain. Specifically, I applied a Variational Autoencoder (VAE) deep learning architecture coupled with LASSO regression to identify clinical

and pharmacokinetic covariates influencing tacrolimus exposure in heart transplant patients. This approach allowed for a compact, interpretable latent representation of time concentration profile of tacrolimus drug absorption, enabling personalized immunosuppressive therapy planning. In summary, this thesis presents a set of machine learning pipelines and network-based frameworks specifically designed for data-scarce biomedical contexts, with specific focus on the follow-up of heart transplant patients. Spanning from omics-based classification to covariate modeling and biomarker discovery, the developed methodologies aim to advance precision medicine by enhancing the interpretability of complex biological data and enabling the early detection of adverse events in cardiac transplantation.

Abbreviations

ACR	Acute Cellular Rejection
AMR	Antibody – Mediated Rejection
AI	Artificial Intelligence
ARIMAX	AutoRegressive Integrated Moving Average with exogenous variables
CL	Clearance
CMV	Cytomegalovirus
CVAE	Conditional Variational Autoencoder
DDT	Transport Document
EMB	Endomyocardial Biopsy
EMBC	Engineering in Medicine and Biology Conference
FN	False Negatives
FP	False Positives
GAN	Generative Adversarial Network
GO	Gene Ontology
hsa-	Homo sapiens (prefix for human miRNAs)
IS	Interaction Score
ISHLT	International Society for Heart and Lung Transplantation
KEGG	Kyoto Encyclopedia of Genes and Genomes
KSVM	Kernel Support Vector Machine
LCa	Lung Cancer
LASSO	Least Absolute Shrinkage and Selection Operator
LSBU	London South Bank University
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
miRNA	microRNA
mRNA	messenger RNA
NSE	Nash–Sutcliffe Efficiency
NTLD	Non-Tumor Lung Disease

OD	Other Diseases
PEA	Pathway Enrichment Analysis
PI	Interaction Probability
PK	Pharmacokinetics
PopPK	Population Pharmacokinetics
RNN	Recurrent Neural Network
RNA	Ribonucleic Acid
SNP	Single Nucleotide Polymorphism
TP	True Positives
TN	True Negatives
VAE	Variational Autoencoder
Vc	Central Volume of Distribution
VMC	Mechanical Ventilation Control
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
KL	Kullback–Leibler
SMA	Soil Moisture Accounting
Tc	Time of Concentration
KGE	Kling–Gupta Efficiency
IDF	Intensity–Duration–Frequency
ET	Evapotranspiration
MA	Moving Average
AR	AutoRegressive
PACF	Partial Autocorrelation Function
ADF	Augmented Dickey–Fuller
ICU	Intensive Care Unit
RISC	RNA-induced Silencing Complex
ceRNA	Competing Endogenous RNA
NSE	Nash–Sutcliffe Efficiency
NTLD	Non-Tumor Lung Disease

OD	Other Diseases
PEA	Pathway Enrichment Analysis
PI	Interaction Probability
PK	Pharmacokinetics
PopPK	Population Pharmacokinetics

Project Management Methodology

Throughout the PhD, the execution and coordination of research activities have been managed according to a structured project management approach. This methodology was inspired by the PMBOK Guide (Project Management Body of Knowledge) [1] developed by the Project Management Institute (PMI), one of the most widely recognized standards for project management across industries. The implementation of such a methodology was not only aimed at maintaining control over deadlines and deliverables, but also at enabling strategic planning in the face of multidimensional complexity and a heterogeneous research context. These methodological skills were also formally developed through targeted training programs and workshops organized by both the host company, I4 Consulting S.R.L., and the academic institution. These courses provided foundational and practical knowledge in project planning, risk management, resource allocation, and the use of collaborative tools.

Overview of the Workflow and Ticket Lifecycle

Each research task, subproject, or development effort was formalized as a ticket and tracked within Jira, a collaborative software tool designed for agile project management. This structured approach

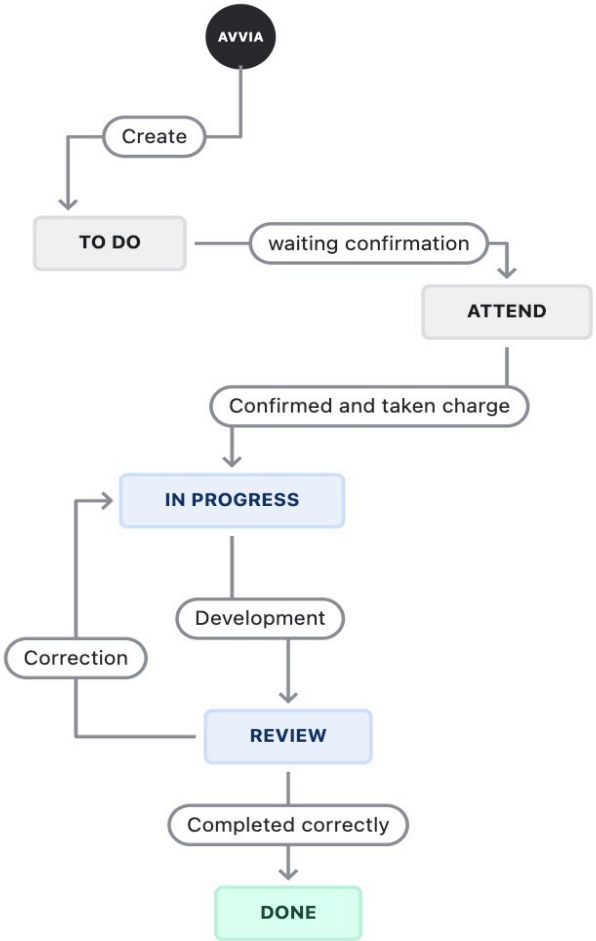


Figure 1. Visual representation of the task management process used throughout the research activities.

allowed for the systematic breakdown of complex research objectives into manageable units, ensuring clarity in task ownership, scope definition, and progress monitoring. The adoption of a formal ticketing system was particularly valuable dealing with multiple projects. The ticket lifecycle followed a well-defined process inspired by PMBOK's process groups (Initiating, Planning, Executing, Monitoring & Controlling, and Closing), adapted to the research context. Unlike traditional industrial projects, academic research often faces frequent shifts in priorities and evolving hypotheses. The flexibility of the ticketing system, combined with a standardized lifecycle, made it possible to respond efficiently to such changes while maintaining methodological coherence. As shown in Figure 1, every task began with the "Create" phase, leading to the "TO DO" status. Tasks were then moved to a temporary "waiting confirmation" status, where responsibilities and scopes were discussed. Once confirmed and formally assigned ("Confirmed and taken charge"), the task entered the "IN PROGRESS" phase, corresponding to the development or experimental implementation. Upon completion of the initial development or analysis, tasks were moved to the 'REVIEW' phase, where they were assessed for correctness and meaningfulness. When needed, corrections were applied iteratively until the task was "Completed correctly", and then officially marked as "DONE". In addition, this approach fostered a culture of accountability and continuous communication between academic and industrial partners. Every status transition was supported by documentation and shared decision-making, reinforcing transparency and facilitating asynchronous collaboration. To complement this operational workflow, Confluence was used as the central platform for documenting all relevant project information, including design choices, meeting reports, protocols, and methodological decisions. The integration of Jira and Confluence established a seamless connection between task execution and knowledge management. This digital infrastructure not only supported the day-to-day operations of the research but also enabled retrospective analyses, audit trails, and cumulative learning throughout the doctoral project. This project management framework served as both a practical tool and a learning environment, shaping a professional mindset rooted in planning, traceability, and adaptability skills that are increasingly essential in data-driven scientific research and collaborative innovation contexts.

Planning and Execution Using Gantt Charts

The implementation of a structured methodology was further supported by Gantt charts, as illustrated in Figure 2. Gantt timelines were automatically generated within Jira and grouped by project and subproject, enabling effective visual tracking of deadlines, dependencies, and progress over time. Each high-level project (e.g., time series modelling, transcriptomic classification, pharmacokinetics analysis) was decomposed into actionable tasks, allowing precise monitoring of deliverables, including publications, implementations, conference and presentations. This approach supported the prioritization of critical paths, identification of inter-task dependencies, and resourceallocation over the course of the PhD. Furthermore, it enabled seamless collaboration between academic supervisors, industry partners, and research teams.

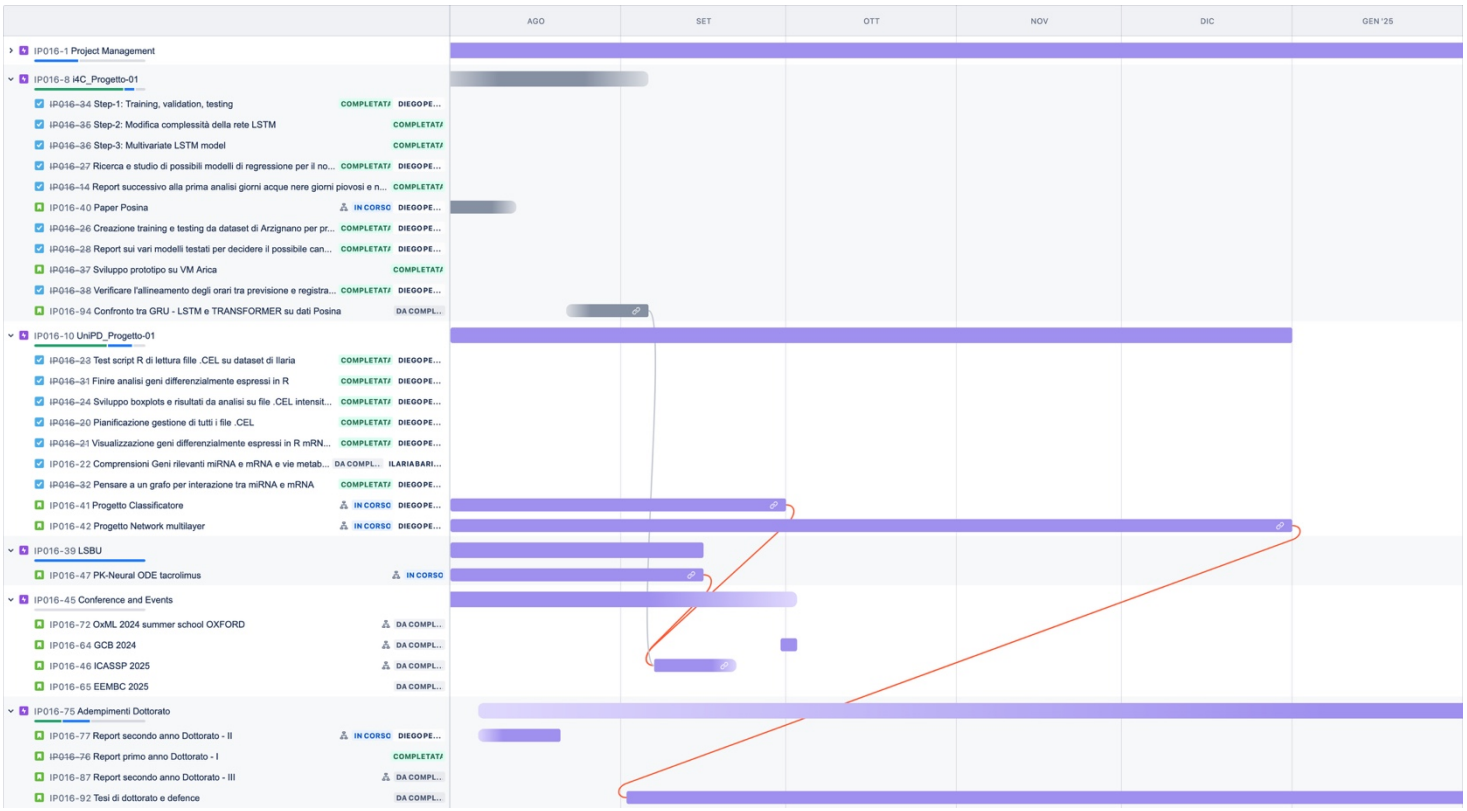


Figure 2. Example of Gantt chart used for planning and monitoring research activities across multiple projects.

Adopting a standardized project management methodology proved essential in addressing the inherent complexity and multidisciplinarity of the PhD journey. This structured approach facilitated more effective coordination between academic and industrial collaborators, enhancing communication and alignment across all phases of the research. It enabled the early identification of potential delays or bottlenecks, thereby allowing for timely and proactive corrective actions. Moreover, it ensured full traceability of each task from initial ideation to final implementation encompassing the management of code, results, and key decisions. The methodology also demonstrated strong scalability, supporting the concurrent handling of multiple research streams with varying timelines and objectives. The integration of project management practices into the research workflow played a pivotal role in ensuring the timely achievement of results and maintaining consistency and coherence throughout the doctoral work.

1 Introduction

1.1 Computational Methods in Biomedicine and Heart Transplantation

In the last decades, the convergence of computational science and biomedicine has catalyzed a transformative shift in how complex biological systems are studied, interpreted, and translated into clinical decision-making. The growing adoption of machine learning and artificial intelligence across diverse medical and biological domains has profoundly impacted research methodologies and healthcare delivery. These technologies have reshaped diagnostic workflows, enabled early disease detection, and facilitated personalized treatment strategies, thereby redefining the role of data-driven intelligence in modern medicine [2,3]. Precision medicine is one of the recent and powerful developments in medical care, which has the potential to improve the traditional symptom-driven practice of medicine, allowing earlier interventions using advanced diagnostics and tailoring better and economically personalized treatments. Intelligent big data platforms are necessary to improve the quality and transition of healthcare by expediting investigation of active hidden factors in clinical data with machine learning algorithms to obtain actionable gap-based information about patients for early detection and prevention. This evolution is largely driven by the increasing availability of high-dimensional biomedical data and the concurrent maturation of machine learning (ML) and deep learning (DL) methodologies, which have demonstrated exceptional performance in uncovering subtle patterns, making predictions, and enabling decision support in highly variable and nonlinear systems [4]. Originally developed within the domains of computer science and statistics, ML algorithms have progressively been adopted across a variety of biomedical applications, ranging from diagnostic imaging and electronic health records analysis to the integration of high-throughput omics data. In particular, the ability of these models to handle data heterogeneity, missing values, and temporal dynamics makes them especially suitable for clinical environments where data are often fragmented and incomplete. Deep learning, a subset of ML inspired by the layered architecture of the human brain, has further extended these capabilities. Architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and variational autoencoders (VAEs) have been successfully applied to learn hierarchical representations of medical images (as histopathological images), physiological time series, and molecular profiles, respectively [5]. These models have moved the field beyond simple classification and regression tasks, towards generative modelling, multimodal integration, and personalized risk prediction. Moreover, ML and DL approaches are playing an increasingly pivotal role in translational medicine, bridging the gap between basic biological discoveries and patient-centered clinical outcomes. In the specific context of heart transplantation, where graft survival and patient prognosis are tightly linked to early and accurate identification of complications, computational methods offer an opportunity to enhance clinical decision-making. Endomyocardial biopsy (EMB) remains the gold standard for rejection diagnosis. In parallel, as observed in fields like oncology [6], molecular and computational strategies are increasingly being investigated to support traditional diagnostics and enable more personalized, patient-tailored immunosuppressive management. To address these goals, the mentioned computational approaches have emerged as promising tools to complement traditional histopathological analysis. These methods can process complex, high-dimensional data derived from endomyocardial biopsies, including transcriptomic, proteomic, and histological image data, and uncover patterns that may not be visible through conventional diagnostic workflows. In recent years, several studies have demonstrated the

feasibility of applying ML and DL models to improve risk stratification or predict acute rejection events. ML algorithms have been employed to analyze gene expression data from EMBs and accurately classify rejection phenotypes [7]. Similarly, multiple deep learning architectures have been used on digitized biopsy slides to detect rejection features with performance comparable to expert pathologists [8]. Other studies have leveraged ML models to predict long-term survival or immunosuppressive therapy response using longitudinal clinical data [9] [10].

1.2 Clinical and Molecular Challenges in Heart Transplantation

1.2.1 The Heart Transplant

The field of cardiac transplantation has undergone significant clinical and technological advancements. Despite being a complex and resource-intensive intervention, advances in surgical techniques, immunosuppressive regimens, and perioperative care have progressively improved post-transplant outcomes. According to the latest report by the International Society for Heart and Lung Transplantation (ISHLT), over 6,000 heart transplants are performed annually worldwide, with ischemic and non-ischemic cardiomyopathies representing the most frequent indications for transplantation [11]. Over the last decade, a progressive shift in the demographics of transplant recipients has been observed. In the last decades, the median recipient age and BMI have increased, as has the proportion of female and allosensitized patients undergoing transplantation [12] [13]. These trends likely reflect both expanded eligibility criteria and improvements in the pre- and post-transplant immunologic management of sensitized individuals. Post-transplant survival is influenced by a variety of clinical variables, including underlying diagnosis, recipient and donor age, sex, need for pre-transplant mechanical circulatory support, such as extracorporeal membrane oxygenation (ECMO), and re-transplantation status. While one-year survival rates are highest among patients with ischemic and non-ischemic cardiomyopathies, individuals with congenital heart disease show significantly reduced long-term survival. Male recipients, older donors or recipients, and patients bridged to transplant with ECMO are associated with worse outcomes [14]. Nevertheless, median survival following adult heart transplantation continues to improve, now reaching 12.5 years, and extending to 14.8 years among those who survive the first-year post-transplant. The principal causes of death in transplant recipients have remained relatively stable over the years and include primary graft failure, rejection, cardiac allograft vasculopathy (CAV), infectious complications, malignancy, renal failure, and multi-organ dysfunction [15]. These outcomes underscore the need for continuous surveillance strategies and improved diagnostic tools to detect early signs of complications and guide timely therapeutic interventions. In this context, the increasing availability of high-resolution and multimodal data, ranging from clinical records and laboratory results to histopathological images and molecular profiles, offers new opportunities to refine patient monitoring [16] [17].

1.2.2 Endomyocardial Biopsies: The Clinical Gold Standard

In heart transplant recipients, cardiac denervation often renders early signs of graft rejection clinically silent, delaying diagnosis until damage becomes more extensive and potentially irreversible. Close surveillance of the transplanted heart is essential to reduce the risk of graft failure, morbidity, and mortality. Endomyocardial biopsy (EMB) remains the gold standard for monitoring cardiac allograft rejection and is routinely used in post-transplant surveillance protocols. The procedure is typically performed via percutaneous access through the internal jugular or femoral vein, guided by fluoroscopy or echocardiography, and involves sampling at least three myocardial tissue fragments from the right ventricle. These samples are then processed and graded histologically according to the ISHLT guidelines, which define standardized criteria for the classification of rejection episodes [18] [19] [20]. EMB frequency is tailored to the patient's immunological risk and the time elapsed since transplantation. Rejection risk peaks within the first 3–6 months post-transplant, which is reflected in standard protocols that recommend weekly EMBs during the first month, biweekly procedures in the second month, monthly until the sixth month, and quarterly thereafter for the remainder of the first year. Additional biopsies are performed as clinically indicated, for example, following a confirmed rejection episode or when clinical suspicion arises. EMB plays a central role in the management of heart transplant recipients and remains the definitive standard for the diagnosis of cardiac allograft rejection. Its integration into surveillance protocols has significantly improved post-transplant outcomes. Nevertheless, as clinical research advances, there is increasing interest in developing complementary tools that can support EMB by providing additional molecular-level insights. Computational and transcriptomic approaches may help capture subtle immunological changes occurring in the graft, even when histological findings are still within normal limits [21]. As a result, there is growing interest in computational approaches capable of extracting quantitative molecular information from EMB samples, to complement or enhance traditional histological evaluation [22,23].

1.2.3 Infection and Rejection in Heart Transplantation

In the first year following heart transplantation, two of the most critical complications affecting patient outcomes are acute cellular rejection (ACR) and cytomegalovirus (CMV) infection. Both conditions can elicit similar histopathological and inflammatory responses in the graft, making their differentiation particularly challenging using conventional diagnostic approaches.

Acute Cellular Rejection (ACR)

ACR remains the most prevalent form of allograft rejection during the first year after transplantation. It results from a mismatch between the donor and recipient major histocompatibility complex (MHC), which activates recipient T lymphocytes through direct or indirect antigen recognition pathways. This immune response triggers a cascade of cytokine release, recruitment of additional inflammatory cells, and ultimately, cytotoxic injury to the myocardial tissue [24]. At the histological level, ACR is characterized by an inflammatory infiltrate predominantly comprised of lymphocytes, macrophages, and occasional eosinophils, associated with myocyte damage. Usually, myocyte injury is accompanied by encroachment of inflammatory cells surrounding myocytes, that show irregular or scalloped borders. Additionally, myocytes can be partially or wholly replaced,

inducing a distortion of the normal myocardial architecture. In 2004, the International Society for Heart and Lung Transplantation (ISHLT) introduced a revised classification system for acute cellular rejection (ACR) to replace the more fragmented 1990 formulation. The goal was to simplify histological interpretation, improve diagnostic reproducibility, and better align biopsy findings with clinical decision-making [18]. The updated nomenclature reduced the number of categories by consolidating overlapping grades and introduced a four-tiered scale, ranging from 0R (no rejection) to 3R (severe rejection). This revised grading system remains the current clinical standard and is based on the extent and distribution of mononuclear infiltrates, the presence and severity of myocyte injury, and the detection of additional features such as edema, hemorrhage, or vasculitis. Figure 3 shows histological examples of the different grades of acute ACR, based on the 2004 ISHLT classification [25]. Below is a detailed description of each grade:

- **Grade 0R – No rejection:** No evidence of mononuclear cell infiltration or myocyte damage. This grade corresponds to "Grade 0" in the 1990 classification.

- **Grade 1R – Mild rejection:** Presence of interstitial and/or perivascular mononuclear infiltrates without distortion of myocardial architecture. Myocyte damage may be absent or limited to a single small focus. This grade merges "Grade 1A" (focal infiltrate without myocyte damage) and "Grade 1B" (diffuse infiltrate with myocyte damage) from the 1990 system.

- **Grade 2R – Moderate rejection:** Two or more discrete foci of mononuclear infiltrates associated with clear evidence of myocyte injury. Eosinophils may be present. This grade approximates the former "Grade 2" (one focus with myocyte damage) and "Grade 3A" (multifocal infiltrates with damage).

- **Grade 3R – Severe rejection:** Dense, diffuse infiltrates of lymphocytes and macrophages, often polymorphic in nature, extensively invade the myocardium. This is associated with widespread myocyte destruction, frequently accompanied by interstitial edema, hemorrhage, and/or vasculitis. It corresponds to the most severe forms of rejection in the 1990 classification [18], including "Grade 3B" and "Grade 4."

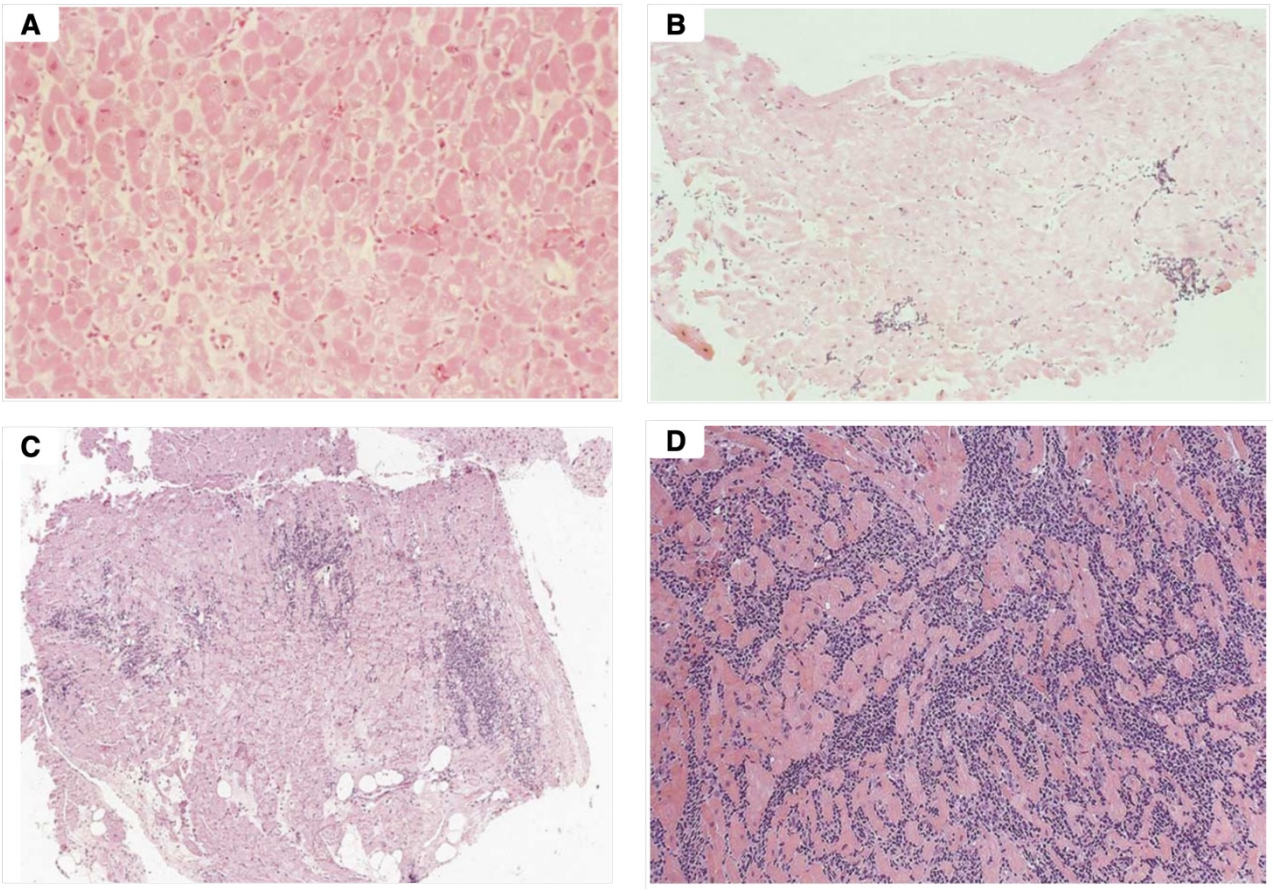


Figure 3. Histological examples of acute cellular rejection (ACR) grades in endomyocardial biopsies, based on the 2004 ISHLT classification. Panel A corresponds to Grade 0R, panel B to Grade 1R, panel C to Grade 2R, and panel D to Grade 3R.

Cytomegalovirus (CMV) Infection

Cytomegalovirus CMV is the most clinically significant viral infection in heart transplant recipients, particularly during the early post-transplant period. As a member of the beta-Herpesviridae family, CMV in the normal population, latently infects 50-90% of individuals, but normally takes an asymptomatic course [26]. This virus has developed several immune escape mechanisms that allow it to actively regulate cellular mechanisms by modulating apoptosis, cytokine activation, leucocyte recruitment, lymphocyte surveillance, and antibody recognition. Thanks to this adaptive strategy, CMV affects innate and adaptive immune control, striking a balance with host clearance mechanisms to persist in the host for life. In transplant recipients, the immunosuppressed state predisposes to both primary CMV infection and viral reactivation, which can originate from either donor or recipient [27]. The cardiac transplanted patients can be stratified according to the serological status of donor (D) and recipient (R) for anti-CMV antibodies. The highest risk is associated with the mismatch between donor and recipient with D+/R- condition. According to the approach chosen by each transplant centre, these patients may undergo pre-emptive therapy (based on the antiviral administration for early asymptomatic CMV viremia detected by surveillance testing) or antiviral prophylaxis (based on the antiviral administration to all patients for the first period after transplantation) [28,29]. At the tissue level, CMV infection can induce inflammatory infiltrates that closely resemble those seen in rejection. Due to this

histopathological overlap between CMV infection and rejection status, histological analysis alone is often insufficient to distinguish between the two. Additional testing such as immunohistochemistry (IHC), polymerase chain reaction (PCR) on blood or formalin-fixed paraffin-embedded (FFPE) tissue is commonly used to detect CMV DNA. However, due to the focal nature of CMV infection and the limited sampling area of endomyocardial biopsies, false negatives are possible. These diagnostic limitations underscore the need for novel biomarkers and integrated analytical approaches that can aid in differentiating infection-driven inflammation from immune-mediated rejection.

1.2.4 MicroRNAs as Biomarkers

MicroRNAs (miRNAs) were first identified in 1993 through genetic screening in Caenorhabditis elegans [30], where they were found to play a critical role in regulating gene expression during development. Since then, miRNAs have been recognized as an essential class of small non-coding RNAs, typically 21 to 25 nucleotides in length, that act as post-transcriptional control by binding to complementary sequences in the 3' untranslated regions (3' UTRs) of target messenger RNAs (mRNAs). Depending on the degree of complementarity, this interaction results in either mRNA degradation or inhibition of translation, thereby modulating a wide range of biological processes Such as pathways involved in immunity, inflammation, cell survival, and tissue remodeling including the ones in cardiovascular system [31,32]. MiRNA biogenesis is a tightly regulated multi-step process. Initially transcribed by RNA polymerase II as a long primary transcript (pri-miRNA), the molecule undergoes nuclear cleavage by the DROSHA-DGCR8 complex, producing a precursor miRNA (pre-miRNA) with a characteristic hairpin structure. This pre-miRNA is exported to the cytoplasm via Exportin-5 and further processed by DICER, generating a mature miRNA duplex. A visual representation of the miRNA biogenesis is depicted in Figure 4[33]. One strand of the duplex is then incorporated into the RNA-induced silencing complex (RISC), where it guides the complex toward specific mRNA targets [34]. In addition to their intracellular regulatory roles, miRNAs have gained considerable attention as potential biomarkers due to their ability to be secreted into the extracellular environment and to remain remarkably stable in biological fluids such as plasma, serum, and urine. This stability is largely attributed to their encapsulation within extracellular vesicles (EVs) or their association with protective carriers like Argonaute-2 proteins and high-density lipoproteins. These features make circulating miRNAs highly attractive candidates for non-invasive biomarker development, particularly in the context of liquid biopsy-based approaches [35]. In transplantation, miRNAs have been studied in EMB tissue samples and liquid biopsy. Intragraft miRNAs raised interest as a companion tool to improve biopsy interpretation. Many studies revealed that miRNA analysis on EMB tissue can discriminate different types of acute rejection with good performances and various platforms.

Nováková and colleagues analyzed the expression level of 11 miRNAs on EMB tissue through Real-Time PCR (RT-qPCR). They delineated an ACR SCORE that distinguishes ACR with a specificity of 91% and a sensitivity of 68% [36]. Furthermore, our group performed a Next-Generation Sequencing analysis for miRNAs on EMB tissue and defined a miRNA signature to discriminate ACR, AMR (Antibody-Mediated Rejection), and MR with a sensitivity of 83.3% and a specificity of 95.8% [37]. On the other hand, Duong Van Huyen demonstrated that miRNA expression is regulated both on graft tissue and on serum. They analyzed miRNA expression on

EMB tissue samples and then assessed their expression level on serum samples matched with EMB. The analysis showed that miR-10a, miR-31, miR-92a, and miR-155 discriminated accurately between patients with and without rejection, with good yield in the external validation cohort[38]. The importance of this work was to directly highlight the link between serum and tissue and, subsequently, it boosted new studies to investigate circulating miRNAs in cardiac allograft rejection. Recently, several groups tried to define a unique miRNA or a miRNA signature associated with acute rejection. Kennel performed a longitudinal profiling of circulating miRNA and showed that hsa-mir-29c-3p and has-mir-486-5p can be potential predictors of ACR and AMR respectively [39], while other groups focused on ACR. Sukma Dewi analyzed serum samples before, during, and after cellular rejection, and showed that hsa-mir-101-3p and hsa-mir-142-3p increased during ACR episodes, with an Area Under the Curve (AUC) of 0.75 and 0.78 [40]. Additionally, in a functional study, she demonstrated that exosomal miR-142-3p is transferred to endothelial cells and compromises endothelial barrier function via down-regulation of RAB11FIP2[41]. A few years later, the research team headed by Crespo-Leiro conducted a prospective, observational, single-center study that included 121 transplanted patients and demonstrated that hsa-mir-181a-5p is significantly overexpressed in patients with ACR compared to no rejecting patients, with a sensitivity and a specificity of 78% and 76% respectively [42]. In 2022 Shah published a multicenter prospective cohort study called Genomic Research Alliance for Transplantation (GRAfT) in which heart transplanted patients from 5 mid-Atlantic centers were phenotyped. This group developed miRNA panels with high diagnostic accuracy for ACR and AMR, and they were validated in an external validation cohort, reaching very high levels of sensitivity (86% for ACR and 82% for AMR) and specificity (76% for ACR and 84% for AMR). Taken together all these data suggest that miRNAs represent a unique class of omics-derived biomarkers, offering both diagnostic and mechanistic insight into allograft biology. The clinical application of miRNA-based biomarkers remains limited by the lack of cross-study reproducibility, largely attributable to cohort heterogeneity, technical variability, and differences in data processing methodologies. This underscores the need for further research efforts and the adoption of standardized computational pipelines, potentially leveraging machine learning models, to identify robust and generalizable miRNA signatures across diverse patient populations and biological platforms. In particular, multimodal computational frameworks, capable of integrating histological, clinical, and other omics-derived data, may represent a critical step toward next-generation strategies for heart transplant monitoring and personalized immunosuppression management.

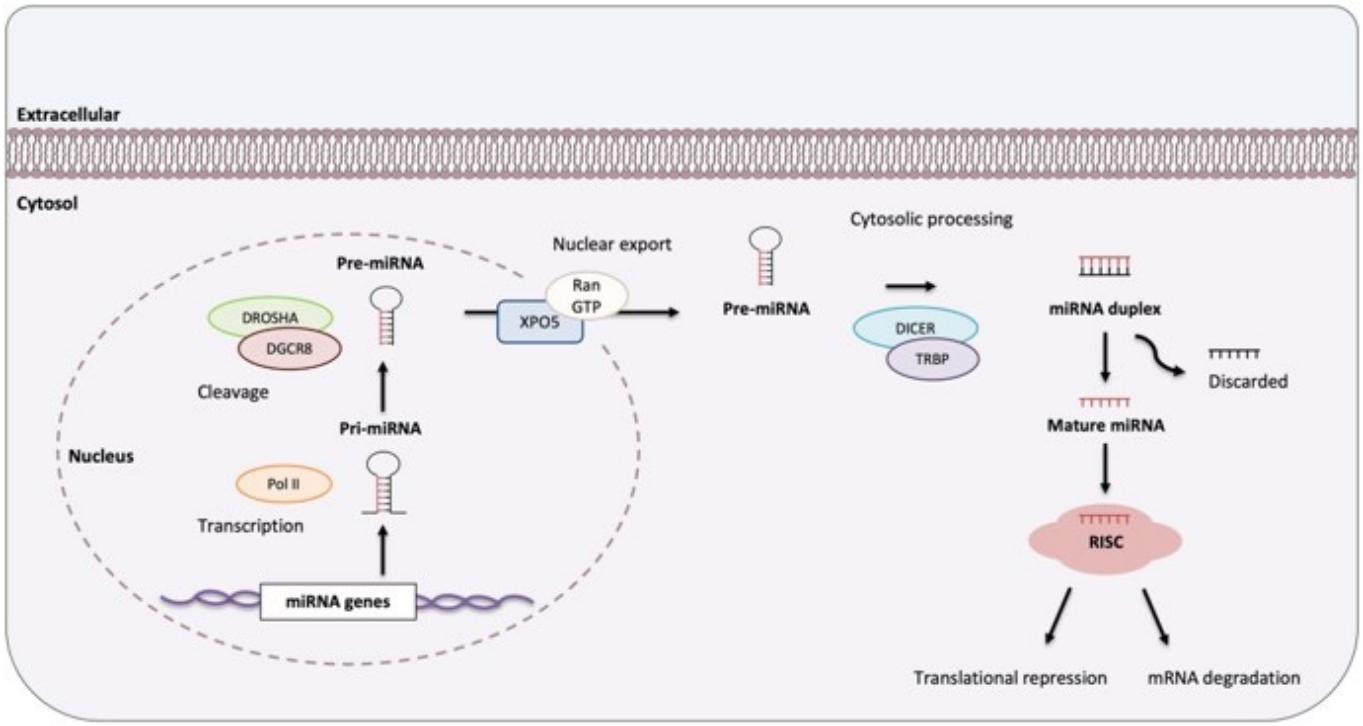


Figure 4. MicroRNA biogenesis: Transcribed by RNA Polymerase II, miRNAs are processed into mature single strands that, once incorporated into RISC, repress translation or degrade target mRNAs

1.2 Machine Learning in Transplantation Research

Since the early days of solid organ transplantation, data collection has followed a highly structured and methodical approach, enabling the development of large observational studies and supporting traditional clinical research. This systematic effort has produced a wealth of data spanning clinical parameters, imaging, histopathology, and patient outcomes. However, the field is now transitioning into a new era, characterized by the availability of increasingly large, heterogeneous, and integrated datasets, including electronic health records (EHRs), medical imaging, and multi-omics profiles, which require more advanced computational methodologies for analysis and interpretation [43]. The field of organ transplantation is undergoing a paradigm shift with the integration of artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), into clinical and research workflows. ML is a subset of artificial intelligence that focuses on the development of statistical models and algorithms capable of learning patterns from data without being explicitly programmed for specific tasks, simulating the human interpretation and generalizability behavior. Formally, ML involves the estimation of a function $f: X \rightarrow Y$ that maps input variables X to output labels Y by minimizing a loss function over a set of training examples. The learning process is highly dependent on the nature of the data and the feedback signal available during training. To help understand when ML might be applicable to solving clinical questions, it can be helpful to consider the different main paradigms [44]:

- **Supervised Learning**: In supervised learning, the algorithm is trained on a dataset that contains input-output pairs. The model learns to map inputs to corresponding outputs by minimizing a predefined loss function [45]. This paradigm is widely used for classification and regression tasks. Supervised learning relies on large, labeled datasets and is one of the most extensively applied approaches in biomedical domains.

- **Unsupervised Learning**: This approach deals with datasets that lack explicit labels. The algorithm attempts to uncover hidden patterns, such as clusters, latent structures, or correlations within the data. Common techniques include k-means clustering, principal component analysis (PCA), and autoencoders [46].

- **Semi-Supervised Learning**: Positioned between supervised and unsupervised learning, semi-supervised learning leverages a small quantity of labeled data alongside a larger pool of unlabeled data. This is particularly valuable in biomedical contexts where labeled data is costly or time-consuming to obtain. Semi-supervised techniques can significantly improve performance in scenarios with data scarcity [47].

- **Reinforcement Learning**: In this paradigm, an agent learns to take actions in an environment to maximize a cumulative reward signal. Rather than learning from labeled examples, the model iteratively interacts with the environment, receiving feedback in the form of rewards or penalties [48].

Deep learning (DL) is a subfield of machine learning characterized by the use of artificial neural networks (ANNs) with many layers (hence “deep”) to model complex, non-linear relationships within high-dimensional data. These models are capable of automatically learning hierarchical feature representations from raw data, thereby reducing the need for manual feature engineering. This characteristic makes DL especially powerful in biomedical contexts where data are often unstructured, heterogeneous, and noisy [49]. Among the core architectures in deep learning, multilayer perceptrons (MLPs) form the foundational structure. An MLP consists of an input layer, one or more hidden layers of artificial neurons (also called units or nodes), and an output layer. Each neuron in a given layer is fully connected to the neurons in the subsequent layer, and applies a nonlinear activation function (e.g., ReLU, sigmoid, or tanh) to a weighted sum of its inputs. The capacity of MLPs to approximate any continuous function over a compact domain is guaranteed by the universal approximation theorem, provided they have sufficient width and depth [50].
While effective on structured data, MLPs are limited in capturing spatial or sequential patterns due to their lack of architectural biases. Convolutional neural networks (CNNs), by contrast, are specifically designed to exploit the spatial structure of image data. By using local receptive fields, weight sharing, and pooling, CNNs efficiently detect features at multiple levels of abstraction, making them the architecture of choice for tasks such as medical image classification and histopathology-based diagnostic models [51]. For sequential or temporal data, recurrent neural networks (RNNs) and their more advanced variants, such as long short-term memory (LSTM)

networks and gated recurrent units (GRUs), are particularly suitable. These networks maintain an internal state that captures dependencies across time, allowing them to process sequences of variable lengths. In clinical settings, they have been applied to model longitudinal patient records, immunosuppressive drug response trajectories, and electrocardiographic signals [52]. In recent years, transformer architectures have redefined the state-of-the-art across multiple domains. Originally developed for natural language processing, transformers employ self-attention mechanisms that allow the model to dynamically weigh the importance of each input element relative to others [53]. This ability to model long-range dependencies without recurrence has led to their adoption in multimodal data integration tasks in biomedicine, where genomic, histological, and clinical information must be jointly analyzed [54].

In transplantation, where timely and accurate assessment of graft health is essential, ML and DL models offer the potential to enhance diagnostic precision, enable early intervention, and support individualized patient management. These approaches have been applied to a wide range of tasks, including prediction of graft survival, detection of rejection episodes, stratification of risk profiles, and identification of novel molecular or imaging-based biomarkers.Liou et al. (2025) evaluated survival machine learning models, such as Random Survival Forests and Cox-based Gradient Boosted Trees, on a large-scale cohort of adult heart transplant recipients from the Scientific Registry of Transplant Recipients (SRTR). The goal was to predict 1-year mortality following heart transplantation, with models achieving moderate predictive performance (C-index around 0.62–0.63) using only pre-transplant clinical features [55]. In the histopathological domain, Lipkova et al. (2022) developed CRANE (Cardiac Rejection Assessment Neural Estimator), a weakly-supervised deep learning framework that automatically detects, subtypes, and grades cardiac allograft rejection from hematoxylin and eosin (H&E)-stained endomyocardial biopsies (EMBs). CRANE achieved expert-level accuracy across multiple external validation cohorts, without requiring pixel-level annotations [56]. Similarly, Seraphin et al. (2023) trained attention-based convolutional networks to predict both presence and severity (0R, 1R, 2R/3R) of acute cellular rejection from digitized EMB slides, achieving AUROC values above 0.85 in internal testing and robust performance on external datasets [57]. Beyond unimodal approaches, recent efforts are increasingly focusing on multimodal learning combining data from different modalities such as genomics, transcriptomics, histology, and clinical parameters to improve robustness and predictive power [58]. Transformer-based architectures, leveraging self-attention mechanisms, have become central in these multimodal fusion tasks, as they allow modeling of complex dependencies between heterogeneous data types. While still emerging in transplantation, this strategy has shown promise in oncology and cardiology, paving the way for applications where multi-layered integration is essential. As the use of ML and DL becomes more prevalent in high-stakes clinical decision-making, the need for interpretability has also gained importance. Explainable AI (xAI) methods, such as attention maps, Lasso regression features selction, and Shapley values are being increasingly employed to ensure model transparency, facilitate clinical adoption, and uncover biologically meaningful insights. This is especially relevant in histopathology and molecular data analysis, where model-driven feature attribution can guide biomarker discovery and hypothesis generation [59]. These approaches are increasingly used to enhance model transparency, facilitate clinical acceptance, and reveal biologically meaningful patterns. They enable researchers to pinpoint which features (e.g., genes, image regions, or clinical parameters) drive the predictions, helping to bridge the gap between algorithmic decisions and clinical reasoning rather than treating models as opaque “black boxes.” This is especially crucial in fields such as histopathology and transcriptomics, where insights

derived from models can inform the discovery of novel biomarkers and the refinement of diagnostic strategies. In these high-dimensional settings, interpretability becomes not only a desirable property but an essential one, guiding downstream validation and enabling integration into clinical workflows. The combination of high-performing predictive models with interpretable outputs is a central direction for computational medicine. In transplantation, where decisions often involve high stakes and limited data, integrating xAI frameworks can foster both trust in model outputs and ensure the biological relevance of their findings [60]. In conclusion, ML and DL are rapidly transforming the field of transplantation by offering data-driven tools capable of handling the complexity and heterogeneity of modern biomedical data. While many challenges remain, such as generalization across centers, data scarcity, and clinical validation, the integration of multimodal fusion techniques and xAI frameworks holds great promise for improving the accuracy, interpretability, and clinical relevance of computational models in transplant medicine.

1.4 Complex Systems and Network Theory in Biology

Biological systems are characterized by a vast number of interacting components genes, proteins, RNAs, metabolites whose relationships define the functional organization of life. Modeling such interactions is crucial for understanding the mechanisms underlying physiological and pathological states. Over the past decades, complex network theory has established itself as a foundational framework for representing and analyzing biological systems, laying the theoretical groundwork for modern systems biology. In this approach, biological components, such as genes, proteins, RNA or even pathways and organs, are represented as nodes, while their interactions whether physical, regulatory, or functional are modeled as edges [61]. This abstraction enables researchers to explore fundamental topological properties, identify central regulatory elements (hubs), detect modular organizations (communities), and infer dynamical behaviors of the system under perturbations. Early applications in biology, such as gene co-expression or protein–protein interaction networks, laid the groundwork for what is now a vast and interdisciplinary domain known as network biology [62]. While traditional (monoplex) network models offer valuable insights, they often fail to capture the multidimensional nature of biological complexity. In real systems, interactions are not uniform but occur across different types and contexts: for instance, transcriptional regulation, post-transcriptional modulation, epigenetic modifications, and spatial constraints.
To address this, multilayer networks have been proposed as a natural extension of classical graph models. In a multilayer network, each layer can represent a specific type of interaction or biological condition, while nodes can exist across multiple layers as replica nodes or state nodes. Inter-layer edges capture interdependencies across layers e.g., regulatory effects of small biological molecules (such as microRNAs), or shared nodes between different omic profiles. This richer structure allows for the integration of diverse biological data while preserving their contextual information, overcoming the limitations of naive data aggregation [63]. A graphical overview of this multiscale network paradigm is illustrated in Figure 5, which depicts how biological networks can be structured from the molecular to the social level. Each layer represents a different level of biological organization (molecular, cellular, tissue, social), and edges encode intra- and inter-layer connections, supporting a system-level understanding of disease mechanisms and information

propagation. Once a complex network model is constructed, a wide range of topological and functional metrics can be computed to characterize its structure and extract biologically meaningful insights. These measures enable the exploration of network organization, the identification of critical regulatory components, and the comparison of network properties across different phenotypic or experimental conditions. Among the most commonly used metrics in network biology are:

- **Degree centrality**: quantifies the number of direct connections a node has. Nodes with high degree, often referred to as hubs, typically play central roles in maintaining the structural integrity of the network or orchestrating regulatory cascades [64]. In protein–protein interaction networks, for example, hubs tend to correspond to essential genes or proteins.

- **PageRank centrality**: originally developed by Google to rank web pages, PageRank has been successfully adapted to biological networks [65] [66]. It quantifies the importance of a node based on the importance of its neighbors, under the assumption that connections from influential nodes weigh more. This recursive process makes PageRank particularly suited for identifying key regulators in complex and noisy networks, such as gene regulatory or miRNA–mRNA interaction networks.

- **Betweenness centrality**: measures the extent to which a node lies on the shortest paths between other nodes. Nodes with high betweenness act as bridges or bottlenecks and may control the flow of information. This metric is useful in identifying potential control points or key regulators in signal transduction pathways [67].

- **Closeness centrality**: reflects the average shortest path from a node to all others in the network, indicating how efficiently information can spread from that node. Nodes with high closeness may act as early responders or signal amplifiers in biological networks.

- **Community detection and modularity**: identify clusters of nodes more densely connected to each other than to the rest of the network. These modules often correspond to biological pathways, co-regulated gene groups, or protein complexes [68,69].

- **Robustness and perturbation analysis**: simulate node or edge removal to evaluate the resilience of the network. This is particularly relevant in biological systems for assessing the impact of gene silencing, mutations, or targeted drug therapy [70].

In multilayer networks where nodes are distributed across multiple layers reflecting distinct biological contexts such as omics modalities, tissues, or disease conditions specialized topological descriptors are required to capture both cross-layer interactions and the context-dependent functional role of each node [71]. Most of the classical metrics developed for single-layer networks, such as degree, betweenness, closeness and Page Rank centrality measure, can be generalized to the multilayer context through appropriate mathematical formulations. These generalizations preserve the interpretative power of traditional network descriptors while incorporating the added complexity

of inter-layer relationships. For example, multilayer versions of centrality can capture the role of a node both within each layer and in the global structure, accounting for interdependencies across interaction types or experimental conditions. Similarly, community detection algorithms have been adapted to identify modules that span multiple layers, revealing functionally coherent groups that persist, or shift, their structure across different biological contexts [72]. This extension of monoplex metrics allows for a unified analysis of complex biological systems where interactions are heterogeneous and context-dependent. Together, these measures support a mechanistic and integrative interpretation of high-dimensional biological data. Unlike conventional analytical pipelines that isolate features or perform univariate differential analysis, the network-based approach embeds data in a relational structure, uncovering emergent properties, multi-layer modularity, and non-obvious functional dependencies that would otherwise remain hidden. Several biomedical studies have demonstrated the utility of multilayer and network-based analyses through the application of centrality and modularity measures. Zitnik and Leskovec (2017) developed a novel computational framework, OhmNet, which models tissue-specific protein–protein interaction (PPI) networks using a multilayer graph structure [73]. In this framework, each layer corresponds to a specific human tissue, and the connections across layers are guided by a hierarchical tissue ontology that reflects biological relatedness (e.g., brain subregions or related immune tissues). By incorporating both intra- and inter-layer interactions, the model captures shared and tissue-specific protein functions, highlighting how multilayer representation can enhance interpretability and predictive power in biological complex systems. In complementary direction, the work of De Domenico et al (2015) [68] introduced a framework for analyzing interconnected multilayer networks through the definition of versatility, a centrality measure designed to evaluate node relevance across multiple layers simultaneously. The metric quantifies the node importance accounting for both intra-layer and inter-layer contributions, identifying versatile nodes that act as central regulators across multiple biological conditions. The approach was successfully applied to various real-world systems, including biological networks. Another recent application of multilayer network is MultiCens, a framework developed by Kumar et al. (2023) [74], used to identify genes that mediate cross-tissue signaling. Unlike single-layer approaches, MultiCens distinguishes intra- and inter-tissue connectivity, enabling the prioritization of genes with systemic influence. The method successfully identified known and novel mediators of hormone signaling across human tissues and revealed disrupted inter-regional brain connectivity in Alzheimer's disease, highlighting its utility in both physiological and pathological contexts.It is possible to observe that some of the works cited above intersect complex network models with machine learning or deep learning methods to enhance both predictive capacity and interpretability. Unlike ML or DL models which typically operate on fixed-length, vectorised representations and optimize objective functions (e.g. loss functions), network-based models are intrinsically relational and topological: they encode interactions, structure, dynamics and hierarchical dependencies, and can more naturally incorporate prior biological knowledge. Recent advances in computational biology are exploring hybrid frameworks, such as: Graph Neural Networks (GNNs), Graph Attention Networks (GATs), and multimodal models that fuse molecular networks with clinical, imaging, or multi-omics data to improve performance on classification, prediction, or biomarker discovery tasks. Wysocka et al. (2023) [75] reviewed deep learning models in oncology that integrate biological networks, such as pathway information or protein–protein interaction (PPI) networks, to support feature selection or embed structural biological priors, resulting in models with improved interpretability and biological plausibility. Similarly, Zhang et al. (2021) described applications of Graph Neural Networks

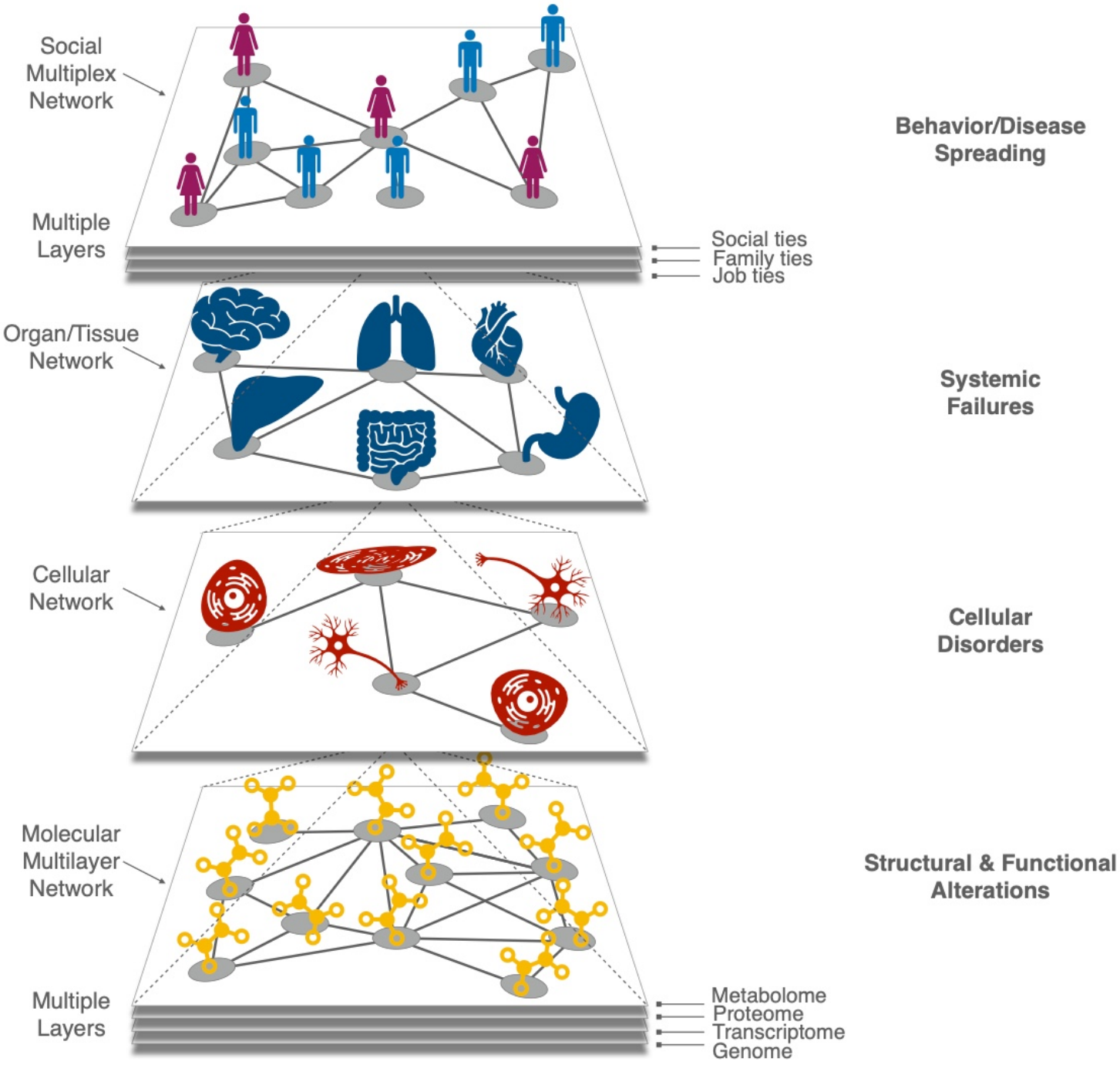


Figure 5. Multiscale network representation from molecular to systemic organ Failures and even disease and social systems. Multilayer networks enable integration across molecular omics (e.g., genome, transcriptome, proteome, metabolome) and support system-level analysis of biological complexity.

(GNNs) to omics datasets, demonstrating that incorporating network topology as part of the input space can enhance disease classification and the prediction of gene or protein function [76]. Although it has not been widely explored yet, in the context of transplantation, complex networks offer multiple applications, particularly in research integrating network analysis with biomarker discovery and their relationships to higher-level representations, such as biological pathways, cellular functions, and organ dysfunctions. Guo et al. (2021) [77] constructed a regulatory miRNA–

mRNA network associated with acute kidney injury after renal transplantation, identifying central miRNAs such as miR-203a-3p and miR-205-5p, and ERBB4 as a hub gene, while pathway enrichment (PI3K-Akt, HIF-1, Ras, MAPK) supported the biological relevance. Similarly, in the study conducted by Lin et al. (2021) they proposed a multi-omics network approach to discover novel miRNA biomarkers (miR-145-5p, miR-155-5p, miR-23b-3p) [78] for prediction and subtyping of kidney transplant rejection, by integrating miRNA-mRNA interactions and protein–protein interaction networks and applying topological metrics like degree, closeness, and betweenness. These applications highlight the potential of complex network models to bridge molecular and clinical information, enabling systems-level interpretations of transplant-related processes. By integrating topological descriptors with functional annotations, expression data and biological knowledge, network-based approaches provide a unique lens to study the regulatory architecture underlying graft injury, immune activation, and tissue remodeling. Complex network architectures are highly transferable to multiple biological contexts, including heart transplantation, where immune-mediated rejection, viral infections, and allograft vasculopathy pose ongoing diagnostic and therapeutic challenges. In particular, multilayer network models offer a structured framework for integrating molecular level data (such as transcriptomic, mirnomic, proteomic, etc.), clinical parameters, and histological scores, allowing for the dissection of complex, multiscale biological phenomena associated with graft dysfunction. Looking ahead, the integration of network theory with machine learning, particularly through the analysis of multimodal data and longitudinal patient monitoring, holds significant promise for improving the follow-up of heart transplant recipients. Such an approach may facilitate the identification of molecular components acting as biomarkers capable of driving phenotypic transitions, thereby enabling the early detection of rejection episodes, the stratification of infection risk, and the prediction of long-term complications. Ultimately, this strategy supports a more personalized and proactive model of patient management. As systems-level methodologies become increasingly accessible, their adoption in clinical transplantation research may help shape the next generation of precision medicine in the field.

2 Omics based Classifier Framework

2.1 Aims

2.1.1 Aim Bootstrap analysis

Given the increasing complexity of omics datasets, a key challenge is not only improving classification performance but also enhancing the transparency and reliability of model decisions. Effective model performance and feature selection are fundamental for explainability and reliability. In many cases, high-dimensional omics datasets suffer from limited number of samples due to clinical constraints, patient conditions, phenotypes rarity and other conditions. Current omics-based classification models often suffer from narrow interpretability, making it difficult to discern meaningful insights where trust and reproducibility are critical. In this study we aim to introduce a machine learning-based classification framework that integrates feature selection with data augmentation techniques to achieve high-standard classification accuracy while ensuring better interpretability. Using the publicly available dataset E-MTAB-8026, we conducted a bootstrap analysis across six binary classification scenarios, with the aim to assess the impact of data availability during training, introduction of synthetic data generation procedures on omics-based classification, accuracy and feature selection of the proposed model. By balancing these factors, our study aims to support the development of robust explainable decision support systems, facilitating omics data exploration and biomarker discovery for personalized treatments. This extensive analysis contributed to the development of a dedicated study, which has been accepted as full contributed paper at the IEEE Engineering in Medicine and Biology Conference (EMBC) 2025 [79].

2.1.2 Aim AugPred pipeline for heart transplant follow-up

Acute Cellular Rejection (ACR) and Cytomegalovirus (CMV) infections represent two of the most significant complications following heart transplantation, each requiring distinct therapeutic strategies to prevent adverse outcomes. Despite the clinical importance of distinguishing between these conditions, their overlapping molecular signatures, coupled with the challenges of data scarcity and variability, have hindered the development of robust diagnostic tools [80–82]. Thus, the primary question this study seeks to address is whether an optimized machine learning pipeline can reliably identify ACR and CMV infections in heart transplant patients, using limited and highly variable microarray datasets. To this end, we introduce AugPred, a machine learning pipeline designed to handle microarray data in post-transplant surveillance. AugPred integrates data augmentation to synthetically expand training sets, L1-regularized logistic regression for interpretable feature selection, and kernel-based Support Vector Machines for classification. In order to validate the robustness of AugPred, we have tested it on a large, publicly available dataset of miRNA profiles from patients with lung cancer and non-tumor lung diseases, by simulating the training of the pipeline on data-scarce scenarios and assessing the generalization performance on the remaining data. This work fills a critical gap in the literature, where few studies have successfully integrated microarray data with machine learning techniques to develop clinically applicable diagnostic tools in data-scarce scenario [83]. While previous efforts have focused on

either improving classification accuracy or enhancing biomarker discovery, AugPred uniquely combines these objectives into a unified pipeline.

2.2 Materials and Methods

2.2.1 E-MATB 8026 dataset description and preprocessing

For bootstrap analysis and validating AugPred pipeline we used a publicly available dataset comprising microarray expression profiles derived from blood samples of over 3,000 individuals. The dataset is available on the public repository of functional genomics data ArrayExpress with accession code: E-MATB-8026 (Link to the dataset). It is related to a study aimed to evaluate the potential of circulating miRNAs as non-invasive biomarkers for detecting lung cancer (LCa) [84]. The samples are categorized into four classes: patients diagnosed with LCa (600 samples), those with non-tumor lung diseases (NTLD, 593 samples), individuals with other diseases (OD, 977 samples), and healthy controls (964 samples). Total RNA was extracted using the Qiagen PAXgene Blood miRNA Kit, labeled and hybridized using the Agilent miRNA Complete Labeling and Hyb Kit, and scanned with the Agilent microarray scanner system to generate miRNA expression profiles.

Preprocessing for bootstrap analysis

miRNAs with invalid expression values were removed to reduce noise and improve data quality. Only miRNAs with "hsa-" prefix were retained, ensuring that the dataset exclusively contained human-associated miRNAs. After this filtering steps, the resulting dataset included the expression of 1,184 miRNAs as features. To create a well-balanced dataset while reducing computational complexity, we selected an equal number of 500 samples per class (LCa, NTLD, OD, and Healthy Control). This balanced subset ensures that the classification model is not biased toward any particular group while maintaining sufficient statistical power for analysis.

Preprocessing for AugPred validation

As for the bootstrap analysis, miRNAs with invalid expression values were removed to reduce noise and improve data quality. Only miRNAs with "hsa-" prefix were retained, ensuring that the dataset exclusively contained human-associated miRNAs. To focus the analysis on well characterized conditions and to simulate a data-scarce scenario during the validation of the AugPred pipeline, samples belonging to the OD class were excluded from the dataset. As a result, the remaining data consisted of 606 LCa samples, 593 NTLD samples, and 964 healthy control samples.

2.2.2 Heart transplant data description and preprocessing

Post-heart transplant dataset used for the analysis represents a set of miRNA expression profiles collected from 11 tissue biopsies, 5 from patients experiencing ACR of the allograft, and 6 from patients with CMV infection. A representative example of the histopathological features extracted of CMV infection and ACR is shown in Figure 6 (Barison et al., 2025) [85]. All patients enrolled underwent cardiac transplantation at the Cardio-Surgery Center Gallucci (Department of Cardiac-

Thoracic-Vascular Sciences, and Public Health at the University Hospital of Padua, Italy), and they were monitored during all first year of follow-up post-transplant. All participants were provided with complete information about the study, and all procedures performed were in accordance with the ethics standards of the institutional and/or national research committee and with the 2013 Helsinki Declaration and its later amendments or comparable ethical standards. The local ethical committee approved the study protocol number 0062556.
Sample collection was performed through EMB procedures, during which small endomyocardial tissue fragments were obtained. Patients with biopsies showing inflammatory infiltrates are classified as ACR, while patients with biopsies indicating acute rejection and microbiological evidence of CMV viremia exceeding 1,000 copies/mL are classified as CMV infection. According to the 2004 guidelines of the International Society for Heart and Lung Transplantation (ISHLT) [82], all rejection cases in the dataset were classified as 3A-2R. Among the infection cases, four were classified as 1A-1R, one as 3A-2R, and one as 3B-2R. The microRNA expression profiles were measured using the Thermo-Fisher Gene Chip miRNA 4.0 Assay microarrays following the manufacturer's instructions. A total of 2562 miRNAs expression profiles have been collected. MicroRNA profiles were preprocessed by filtering out those without the 'hsa-' prefix, as they are not associated with the human organism. All missing values were removed from the dataset to prevent the introduction of bias. After the preprocessing step, 2510 miRNAs expression profiles have been retained.

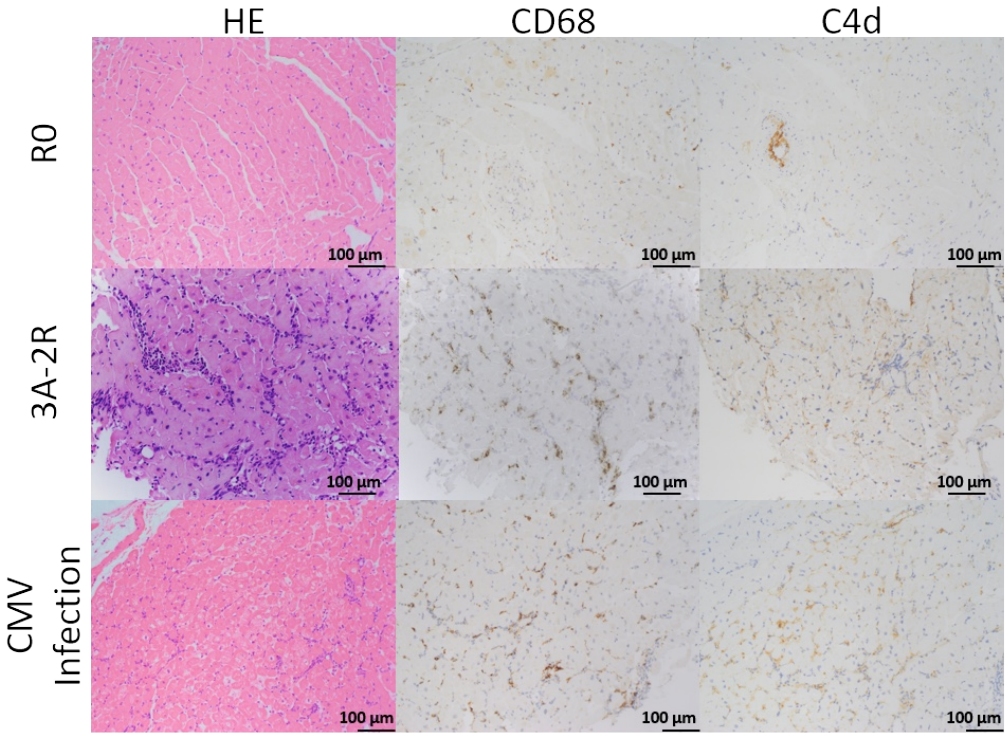


Figure 6. Histological and immunohistochemical analysis of endomyocardial biopsy samples in a control case (R0), a rejection case (3A–2R), and a CMV infection. Samples were stained with Hematoxylin–Eosin (HE), and with anti-CD68 and anti-C4d immunohistochemistry. Scale bar: 100 µm.

2.2.3 Bootstrap analysis model architecture

We designed and studied a model classification framework which consists in the application of 3 main steps.

1) *Synthetic Samples Generation*: To mitigate the cases of data scarcity and enhance model robustness, a Gaussian noise-based augmentation strategy is applied [86] to the training data to generate synthetic samples while preserving the original data distribution. The process involves selecting real samples from each class and perturbing their values with Gaussian noise, where the standard deviation of the noise is set as the 10% of the original feature's standard deviation. For each feature, we independently compute the standard deviation within each class, ensuring that variations reflect intrinsic variability. New synthetic samples are generated by randomly selecting values from the original dataset and adding the computed noise. This method maintains the statistical properties of the dataset while introducing controlled variability to improve model generalization. The final augmented dataset includes an equal number of synthetic samples per class.

2) *Feature Selection*: To tackle high-dimensional datasets, a feature selection approach based on L1-regularized logistic regression (LASSO) is applied. This method penalizes feature coefficients, effectively reducing the weights of irrelevant features to zero, thereby selecting only the most influential ones [87]. Multiple LASSO simulations are conducted; each trained on a dataset augmented with synthetic samples. Inverse regularization strength parameter (c) has been set to 0.01. After each simulation, features with non-zero L1 coefficients are recorded. The occurrence of each selected feature across multiple simulations is counted, and only those present in more than 50% of the executed simulations are retained and used as input features for the subsequent classification step.

3) Kernel-Based Classifier: Kernel Support Vector Machine (KSVM) model with polynomial kernel is employed. KSVM is a powerful method that exploit a kernel function to project data into a higher-dimensional space, defining the optimal hyperplane that maximally separates the two classes [88]. The formulation can be denoted from the (1, where α_i are the learned weights, y_i are the class labels, and xi are the support vectors. The function $K(x_i, x)$ is the kernel function that maps data into a higher-dimensional space, and b is the bias term that shifts the decision boundary. n is the number of support vectors and $f(x)$ is the decision function that determines the classification of x. The classifier is trained using the selected features from the feature selection step, ensuring that only the most relevant ones contribute to the decision boundary. Once trained, the KSVM predicts class labels for the original test samples, and performance is evaluated using standard metrics such as accuracy and confusion matrices. For simplicity, we refer to this model architecture as the L1-KSVM Framework.

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b \qquad (1)$$

2.2.4 Bootstrap analysis experimental procedure

To comprehensively evaluate both classification performance and the robustness of feature selection under varying data constraints, a bootstrap-based experimental framework was implemented. Six binary classification tasks were defined from the E-MATB 8026 dataset, by considering all pairwise combinations among the four available classes: LCa, NTLD, OD, and healthy controls. The experimental procedure is designed to evaluate the effect of sample size on classification performance and the interpretability of feature selection. This comprehensive pairwise design enables a systematic assessment of model behavior across a wide spectrum of clinical contrasts, ranging from healthy-vs-disease to inter-disease comparisons with potentially overlapping biological signatures. A complete visual representation of the process is provided in Figure 7. The boostrap analysis have been conducted on the following six binary classification scenarios:

- Scenario 1: Healthy Controls vs LCa
- Scenario 2: Healthy Controls vs NTLD
- Scenario 3: Healthy Controls vs OD
- Scenario 4: LCa vs NTLD
- Scenario 5: LCa vs OD
- Scenario 6: NTLD vs OD

For each predefined classification scenario, we assessed different training set sizes by randomly selecting subsets from the balanced dataset, which consists of 500 samples per class. The training set sizes tested were: 350 (70%), 300 (60%), 250 (50%), 150 (30%), 100 (20%), 25 (5%) and 10 (2%). All remaining samples constituted the test set. Each experiment was repeated 100 times per sample size, with new bootstrap samples drawn in each iteration. These scenarios were selected not only to assess model performance, but also to explore the discriminative power of microRNA profiles under varying degrees of phenotypic similarity and heterogeneity. Three classification approaches were evaluated:

- L1-KSVM Framework incorporating 200 synthetic samples per class, generated using a Gaussian noise-based augmentation technique.

- L1-KSVM Framework without data augmentation, applying the same feature selection and classification procedure but without synthetic data augmentation.

- A baseline LASSO logistic regression model, where features were retained solely based on having a nonzero L1 coefficient.

The trained models were evaluated on the test set, and classification accuracy was recorded. For each sample size and across the six classification scenarios, the number of microRNAs selected by the regularization method was also tracked. By contrasting models with and without data augmentation, we aim to isolate the effect of synthetic variability on the generalization ability and

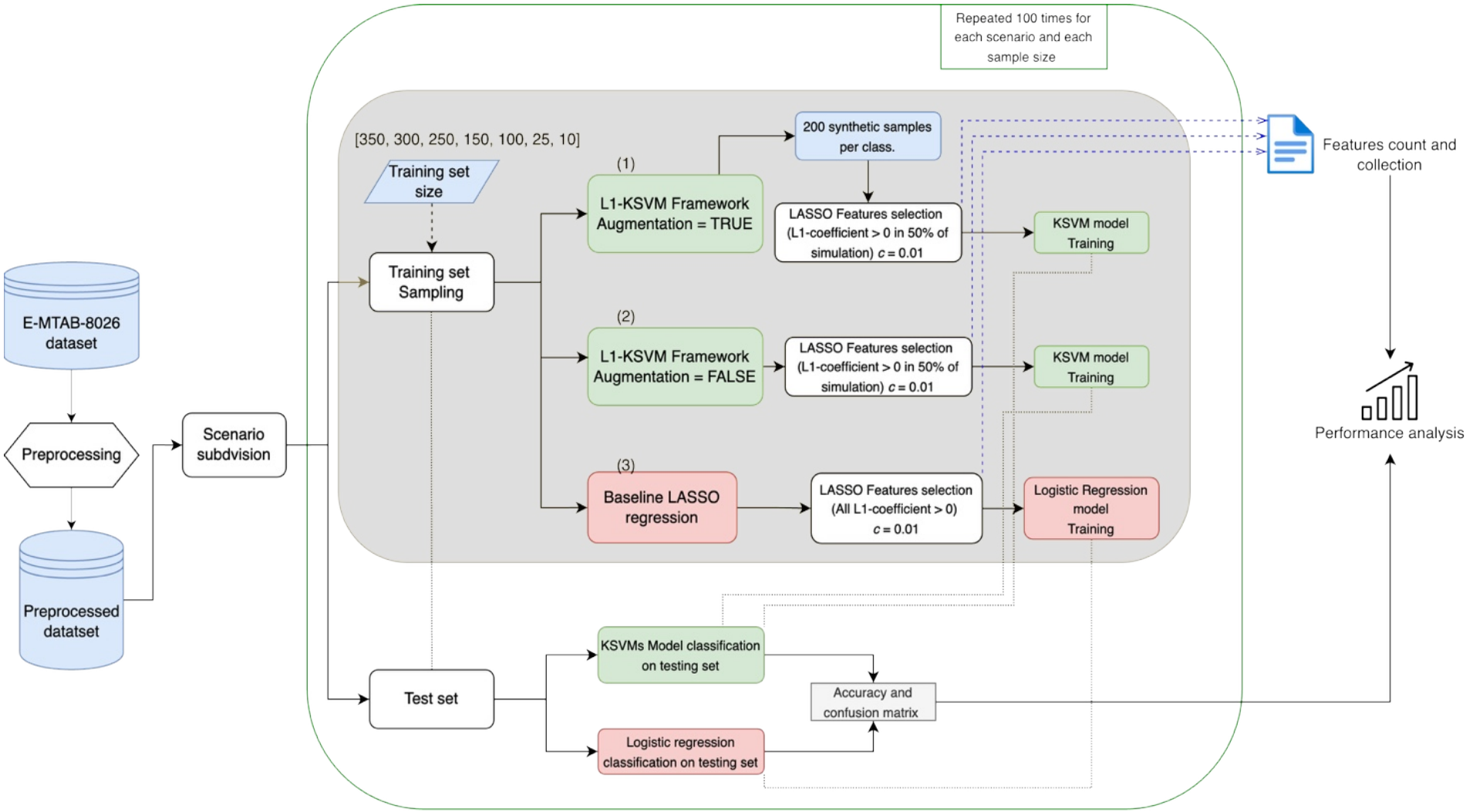


Figure 7. Experimental Procedure for the Bootstrap Analysis. The preprocessed dataset is divided into six binary classification scenarios. For each scenario,training set sampling is performed at different sizes, while the remaining samples constitute the test set. Three classification approaches are evaluated: (1) the L1-KSVM Framework with augmentation, where 200 synthetic samples per class are generated using a Gaussian noise-based augmentation technique, (2) the L1-KSVM Framework without augmentation, which follows the same feature selection and classification procedure without synthetic data, and (3) the Baseline LASSO regression model, where features are retained if their L1-coefficient is greater than zero. Features collection and classification is then performed. rsulting data are collected to conduct performance analysis.

sparsity of the resulting classifiers. Additionally, including a linear baseline provides a reference point to contextualize the benefits of kernel-based methods.

2.2.5 The AugPred pipeline

Following the model framework studied in the boostrap analysis, we have created the AugPred data-driven pipeline to robustly select relevant miRNA biomarkers and use them to identify ACR versus CMV is composed of multiple steps:

1) Data augmentation
2) Feature selection
3) Classifier training

Figure 8 A) depicts a better visual representation of the system developed.

Step 1: Data augmentation

The first step of the pipeline generates multiple synthetic samples for each class using the training set of our available data. For each specific feature f_i (miRNA expression level), the pipeline we compute the means μ_{ACR} and μ_{CMV}, and standard deviation σ_{ACR} and σ_{CMV} for the two classes. Then, we create 100 new synthetic expression profiles each having $f_{i,synth}$ sampled from the corresponding normal distribution $\mathcal{N}(\mu_{ACR}, \sigma_{ACR})$ or $\mathcal{N}(\mu_{CMV}, \sigma_{CMV})$ Representative results of the augmentation are shown in Figure 8 B).

Step 2: Feature selection

The AugPred pipeline proceeds with a logistic regression model coupled with the L1 lasso regularization, trained on the new augmented dataset. The L1 regularization method, also known as Lasso (Least Absolute Shrinkage and Selection Operator), adds a penalty term (L1 coefficient) to the loss function proportional to the absolute values of the coefficients (weights) of the model, shrinking to 0 features with large enough penalty term and maintaining only non-zero coefficient features [83,89]. The regularization strength can be controlled through a specific parameter. In this study, we utilized the LogisticRegression method implemented in the scikit-learn library [90], which uses the parameter C as the inverse of the regularization strength. Smaller values of C indicate stronger regularization. In this analysis, C was set to 0.01. The generation of the augmented dataset, along with the subsequent feature selection, is repeated 100 times to mitigate the effects of randomness in the sampling process (in Figure 7, this step is represented by the rounded blue line). Features that consistently exhibited a non-zero coefficient in the logistic regression in more than 50% of the 100 repetitions are considered stable.

Step 3 Classifier training

The data in the training fold is then augmented with the same procedure described in Step 1 but considering only the stable features. The augmented dataset is used to train a KSVM model with a polynomial kernel of degree 3 and regularization strength C, set to 1.0, following the default hyperparameter settings [88] [90]. polynomial kernel of degree 3 and regularization strength C, set to 1.0, following the default hyperparameter settings [88] [90].

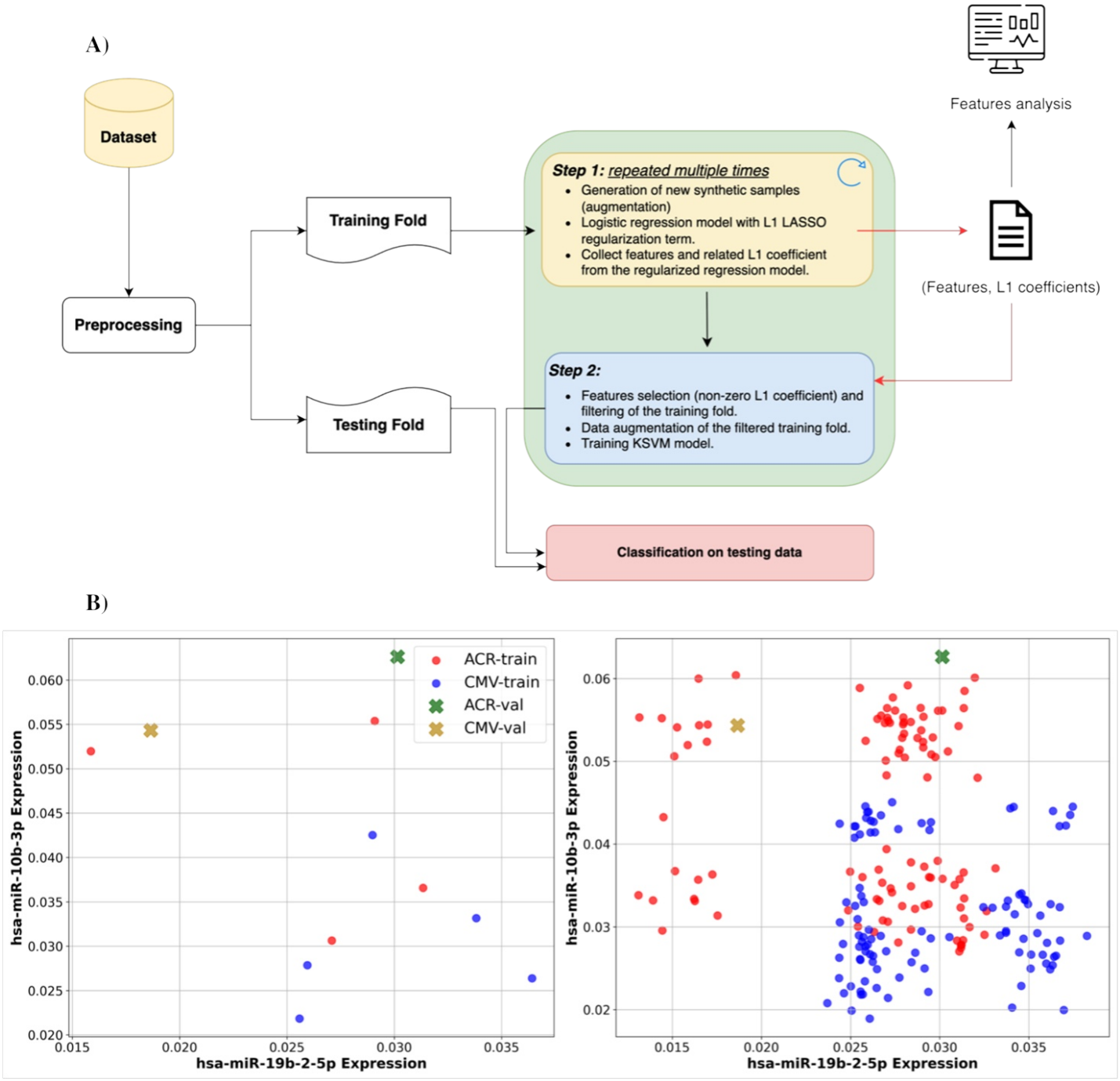


Figure 8. A) Schematic representation of the AugPred pipeline. The dataset is divided into independent training and testing folds, ensuring no data leakage. Data augmentation is applied exclusively to the training fold to improve model generalization. The pipeline consists of two main steps: (1) iterative feature selection using L1-regularized logistic regression on the augmented training fold to identify relevant features and (2) further filtering and augmentation of the selected features, followed by training a KSVM for classification. The testing fold is reserved for final performance evaluation. B) Example visualization of one-fold from the cross-validation process, comparing training fold before and after augmentation. Left: Real measured training samples only. Right: Same fold after data augmentation with synthetic samples. The plot displays the expression values of two randomly selected miRNAs (hsa-miR-19b-2-5p and hsa-miR-10b-3p) as an illustrative example. ACR and CMV samples are color-coded and marker-shaped to indicate their role in training (dots) and validation (crosses) sets.

2.2.6 AugPred Experiment 1: Heart transplant dataset classification

The performance evaluation of the classification was conducted using a 5-fold cross-validation approach to ensure a robust and unbiased assessment of the AugPred pipeline. The pipeline was specifically designed to prevent data leakage [91], which could otherwise lead to performance overestimation or introduce biases. To ensure robustness, steps 1-3 (data augmentation, feature selection, and classification) were conducted within the training fold of each of the cross-validation round, and only the final classifier on the stable features was applied on the test fold. MicroRNAs identified as relevant for classification by the logistic regression model with L1 regularization were further analyzed to assess statistically significant differences in signal intensity distributions between infection and rejection cases. We investigated the pathways in which statistically relevant miRNAs are involved, exploiting the miRNA pathway analysis tool DIANA [92], which integrates experimentally supported interactions and computational predictions, providing a comprehensive overview of miRNA involvement in key biological processes.

2.2.7 AugPred Ablation study on heart transplant data

To evaluate the impact of the different parts of the proposed pipeline, we tested the developed pipeline's (AugPred - KSVM) performance against two alternative approaches. The first is a baseline logistic regression classifier (Baseline approach) that does not incorporate data augmentation or feature selection methods. The second is a simplified version of the pipeline, where the KSVM model is replaced with a logistic regression classifier without a regularization term (AugPred - Logistic).

2.2.8 AugPred Experiment 2: Data scarce scenario simulation on E-MATB 8026 dataset

In order to test the performance of the pipeline when applied to a data-scarce setting as the transplant data, we chose to use a large dataset of miRNA data from LCa patients and healthy control, and simulate the application of the pipeline on a small subset of the available data (data-scarce scenario), and then testing the resulting model on the remaining, much larger, data.

Data scarce scenario
We investigated the classification capabilities in the following 3 binary classification tasks to reflect clinically relevant challenges:

- Healthy control vs LCa
- Healthy control vs NTLD
- LCa vs NTLD

Data-scarce Cross-validation analysis
For each scenario, 80 samples (40 from each class) were randomly selected. The performance AugPred pipeline was assessed using a 4-fold cross-validation (k=4).

Generalization Performance analysis

After the cross-validation performance was obtained, the pipeline was applied to the full of 80 samples (40 per class) and tested on the remaining samples in the dataset, which included 606 LCa, 593 NTLD, and 964 healthy control samples. Performance on the remaining samples was compared to the performance obtained from the data-scarce cross-validation.

Extremely Data scarce scenario

A third analysis was conducted focusing on the LCa versus NTLD scenario to assess the resilience of the AugPred pipeline. We followed the same steps as in the data scarce scenario. For each scenario, 10 samples (5 from each class) were randomly selected from the 606 LCa and 593 NTLD cases, representing approximately 1.6% of the entire dataset. The performance AugPred pipeline was assessed using a 4-fold cross-validation (k=4). To ensure the reliability of the results, the process was repeated 100 times.

2.3 Results

2.3.4 Bootstrap analysis results

Figure 9 illustrates the classification accuracy across six binary classification scenarios for different training set sizes and classification methods. As expected, an increase in training sample size generally improves classification performance. Table 2 rovides a detailed breakdown of performance metrics, including accuracy (Acc. %), true positives (TP %), true negatives (TN %), false positives (FP%), false negatives (FN %) and the accuracy collected from the cross-validation analysis (Cross-Val Acc. %). The Baseline LASSO regression model (Baseline LASSO reg.) achieves relatively higher accuracy in multiple cases, resulting in an overall better performance across the six scenarios. However, Table 1, which summarizes the average number of selected microRNAs (#miRNAs) per training sample size, highlights a key advantage of the L1-KSVM-based approaches, their ability to perform feature selection efficiently. Compared to the Baseline LASSO regression, which retains a substantially larger number of features, both L1-KSVM approaches significantly reduce feature dimensionality, making them more interpretable and easier to investigate. Notably, the L1-KSVM framework with data augmentation improves classification accuracy over its non-augmented counterpart, achieving performance comparable to the Baseline LASSO regression, particularly in scenarios with limited training data. It is also possible to notice the positive effect of data augmentation, particularly evident in scenarios with limited training data, where it enhances model robustness and helps mitigate performance degradation. It is also possible to notice that the accuracy performances of the proposed framework, obtained from the cross-validation analysis of each training sample size, are aligned with the accuracy of the same trained model applied to a larger dataset.

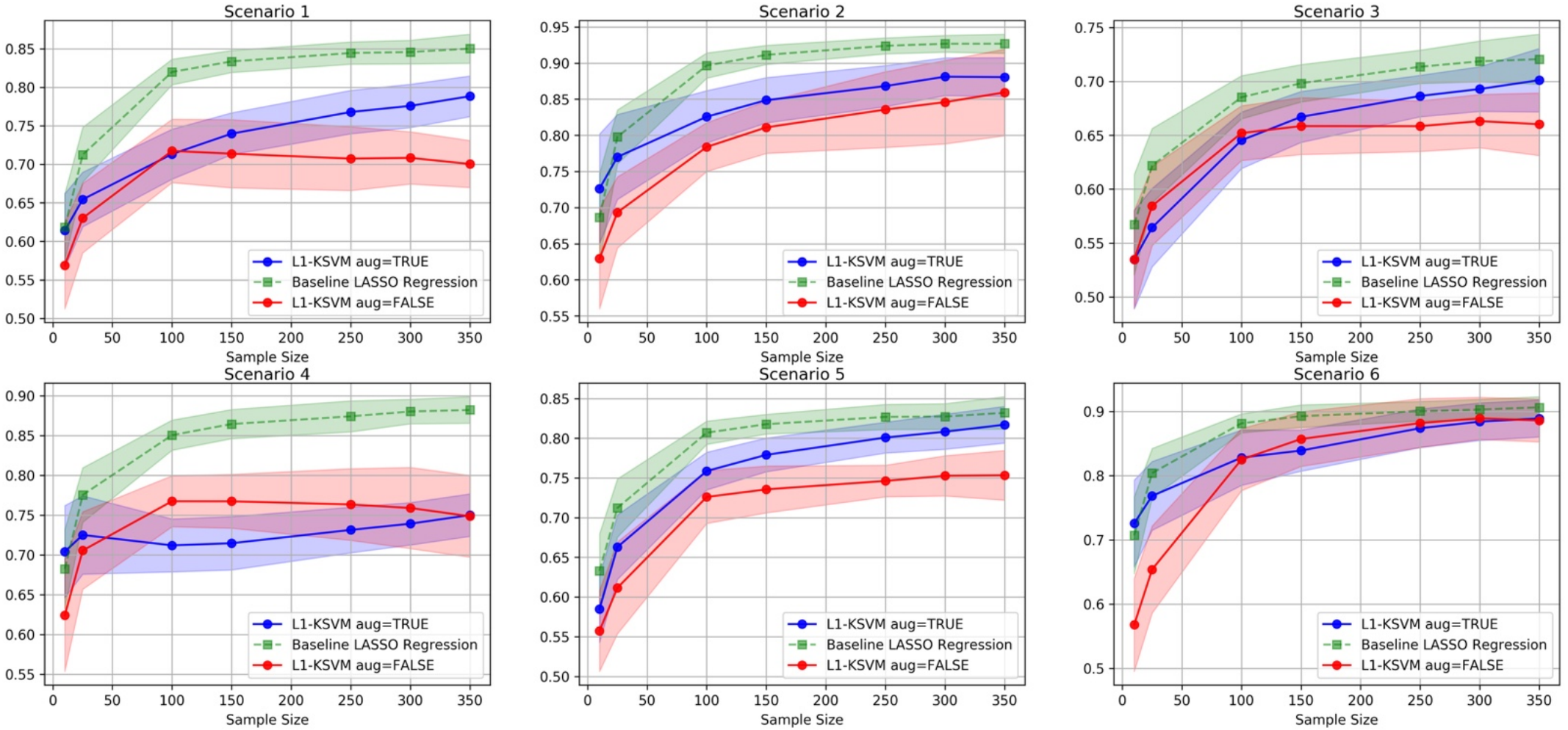


Figure 9. Each subplot corresponds to a binary classification scenario: Scenario 1 (Healthy Controls vs. LCa), Scenario 2 (Healthy Controls vs. NTLD), Scenario 3 (Healthy Controls vs. OD), Scenario 4 (LCa vs. NTLD), Scenario 5 (LCa vs. OD), and Scenario 6 (NTLD vs. OD). The x-axis represents the training sample size, while the y-axis shows the classification accuracy. The plots compare three approaches: the L1-KSVM Framework with Data Augmentation (blue solid line), incorporating 200 synthetic samples per class; the L1 KSVM Framework without Augmentation (red solid line), applying the same feature selection and classification process but without synthetic data; and the Baseline LASSO Regression Model (green dashed line), which retains all features with a nonzero L1 coefficient.Shaded areas indicate standard deviation across 100 bootstrap iterations per sample size.

Training set size	L1-KSVM Aug. (# miRNAs)	L1-KSVM without Aug. (# miRNAs)	Baseline LASSO reg. (# miRNAs)
10	66.54	5.68	1180
25	79.91	10.33	1183
100	96.41	19.93	1183
150	99.56	20.39	1183
250	103.23	20.72	1183
300	104.56	20.72	1183
350	107.05	20.75	1183

Table 1 Average number of selected microRNAs across different classification scenarios

Sample Size	Acc. %	TP%	TN%	FP%	FN%	Cross-val Acc. %
Baseline LASSO regression						
10	64.9	32.74	32.16	17.26	17.84	61.2
25	73.7	37.08	36.64	12.92	13.35	71.7
100	82.3	41.82	40.52	8.17	9.48	81.8
150	83.6	42.60	41.02	7.39	8.97	83.2
250	84.7	43.24	41.47	6.75	8.52	84.6
300	85.0	43.42	41.60	6.57	8.39	83.8
350	85.2	43.54	41.75	6.45	8.24	84.1
L1-KSV (Aug = FALSE)						
10	58.0	29.46	28.58	20.54	21.42	57.1
25	64.6	32.40	32.24	17.60	17.75	62.0
100	74.5	37.11	37.41	12.88	12.58	73.6
150	75.7	37.73	37.98	12.26	12.01	76.3
250	76.5	38.26	38.28	11.74	11.71	76.7
300	76.9	38.25	38.72	11.74	11.27	79.1
350	76.7	38.14	38.64	11.85	11.35	80.0
L1-KSV (Aug = TRUE)						
10	64.8	33.34	31.48	16.65	18.52	66.1
25	69.0	35.03	34.05	14.96	15.94	64.5
100	74.7	36.76	37.95	13.23	12.04	70.4
150	76.4	37.62	38.84	12.37	11.15	72.5
250	78.8	38.86	39.94	11.13	10.05	74.5
300	79.6	39.40	40.28	10.59	9.71	77.2
350	80.4	39.82	40.61	10.17	9.39	79.2

Table 2. Performance metrics across sample sizes for each method

2.3.1 AugPred Experiment 1 Results: Heart transplant dataset classification

Classification performance

The model's performance is validated using cross-validation to evaluate its classification behavior. We tested the AugPred pipeline and perform the cross-validation on all 11 real measured samples. Figure 10 depicts the confusion matrix of the classification process, illustrating the overall performance of the model on the test fold data. The model successfully distinguishes between rejection and infection in nearly all cases, achieving an accuracy level of 91%.

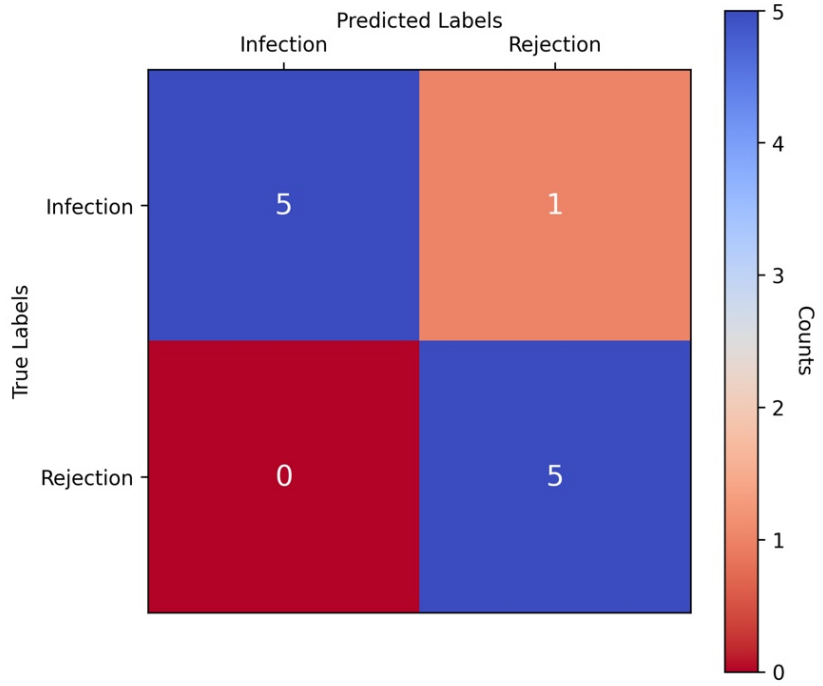


Figure 10. Confusion Matrix of Cross-Validation Classification Results. Confusion matrix showing the performance of the AugPred pipeline in classifying 11 real measured samples into infection and rejection cases using cross-validation.

It is possible to notice that the model has misclassified only one subject, predicting it as a rejection case instead of an infection. This misclassification occurred on the fold number 3 of the cross-validation.

Features and bioinformatic analysis

The L1 regularization in step 1 of the pipeline selected the relevant miRNAs from the augmented generated dataset. Across the different folds, some miRNAs are consistently maintained, while others appear only in specific folds. Table 3 displays all the miRNAs selected in at least two folds of the cross-validation process (column: Fold Count). The table also includes the number of times each miRNA exhibits a non-zero coefficient (columns: Fold 1, Fold 2, Fold 3 and Fold 4) across the 100 iterations.

Name	Fold Count	Fold-1	Fold-2	Fold-3	Fold-4
hsa-miR-30a-5p	4	100	44	100	100
hsa-miR-10b-3p	4	100	99	41	100
hsa-miR-181a-2-3p	3	100	0	93	100
hsa-miR-19b-2-5p	3	100	95	0	93
hsa-miR-486-5p	3	0	98	100	89
hsa-miR-130a-5p	4	55	100	60	68
hsa-miR-514a-3p	3	50	100	0	99
hsa-miR-1307-5p	3	0	99	4	100
hsa-miR-325	3	54	17	0	98
hsa-miR-490-5p	3	30	100	0	28
hsa-miR-105-3p	3	100	0	1	26
hsa-miR-127-5p	3	36	40	0	26
hsa-miR-149-3p	2	0	0	4	89
hsa-miR-30b-5p	2	0	1	0	56
hsa-miR-138-1-3p	2	6	45	0	0
hsa-miR-30a-3p	2	37	0	0	1
hsa-miR-222-5p	2	0	0	5	21
hsa-miR-487b-5p	2	4	0	0	12
hsa-miR-1323	2	12	0	0	1
hsa-miR-198	3	98	96	0	4
hsa-miR-193b-3p	2	0	2	0	1

Table 3. Each microRNA is associated with the number of folds in which it appears (fold count) and the number of times each microRNA exhibits a non-zero coefficient across the 100 iterations.

Name	Fold-1	Fold-2	Fold-3	Fold-4
hsa-miR-30a-5p	100	-	100	100
hsa-miR-10b-3p	100	99	-	100
hsa-miR-181a-2-3p	100	-	93	100
hsa-miR-19b-2-5p	100	95	-	93
hsa-miR-130a-5p	55	100	60	68
hsa-miR-325	54	-	-	98
hsa-miR-105-3p	100	-	-	-
hsa-miR-198	98	96	-	-
hsa-miR-486-5p	-	-	100	89
hsa-miR-514a-3p	-	100	-	99
hsa-miR-1307-5p	-	99	-	100
hsa-miR-490-5p	-	100	-	-
hsa-miR-149-3p	-	-	-	98

Table 4. MicroRNAs that exhibit a non-zero coefficient in more than 50% of the 100 iterations. The '-' indicates either the microRNA was not selected in the fold or did not exceed the 50% threshold.

Table 4 summarizes the miRNAs that display a non-zero coefficient in more than 50% of the 100 iterations for each k-fold. Among these miRNAs, we aim to identify those that exhibit statistically significant differences in signal intensity distributions between the two classes. To this end, we conduct a Kolmogorov-Smirnov test to compare the signal intensity distributions of these miRNAs between the infection and rejection groups. We have collected the miRNAs with a statistical significance difference in Table 5 and depicted the related distribution density plot for each class in Figure 11.

Name	p-value
hsa-miR-514a-3p	0.025974
hsa-miR-198	0.025974
hsa-miR-490-5p	0.047619

Table 5. MicroRNAs that exhibit non-zero coefficient in more than 50% of the 100 iterations and have a statistical different distribution of the signal intensity between the groups.

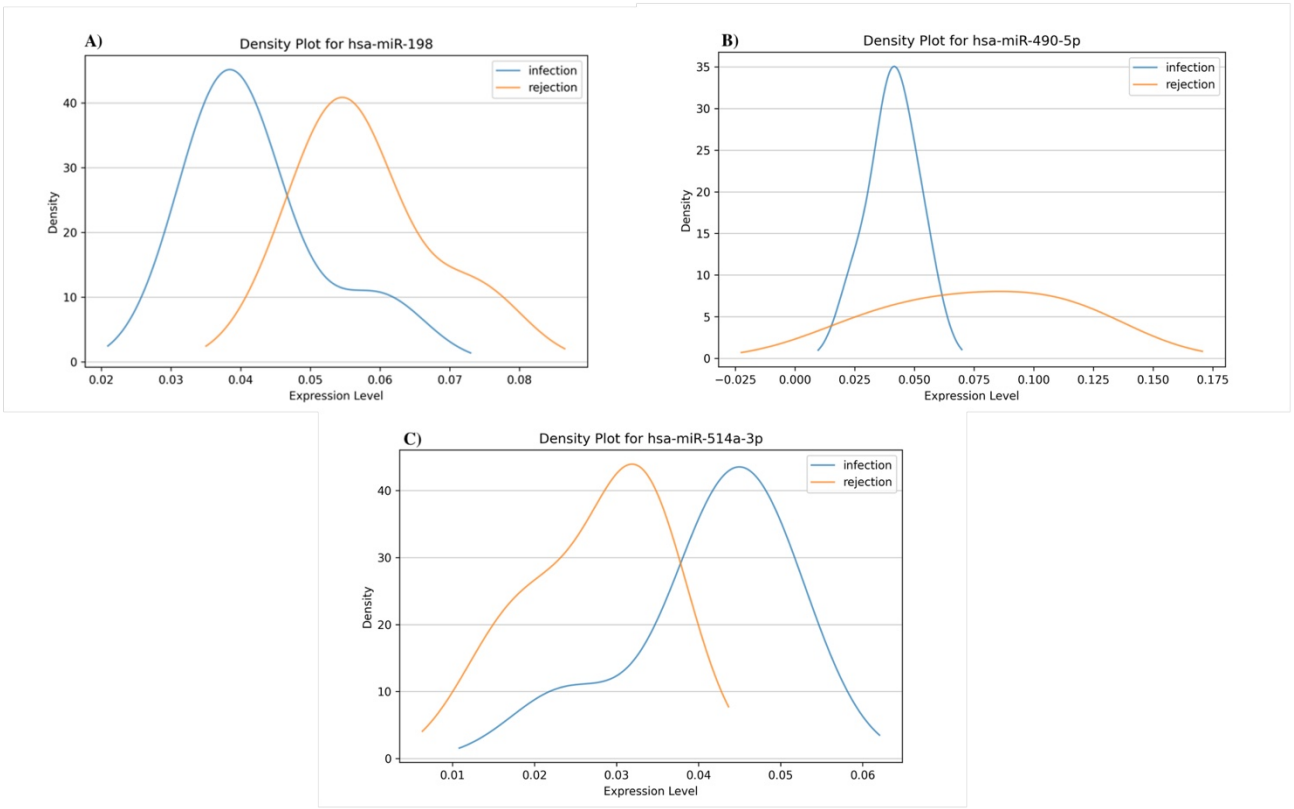


Figure 11. Signal intensity distribution of microRNAs with a non-zero coefficient in over 50% of 100 iterations and statistically significant differences between groups. Density plots illustrating the expression level distributions of the statistically significant microRNAs identified by the Kolmogorov-Smirnov test. (A) hsa-miR-198, (B) hsa-miR-490-5p, and (C) hsa-miR-514a-3p. The blue line represents the infection class, while the orange line represents the rejection class. Differences in the distribution highlight the potential of these microRNAs as distinguishing features for classification between the two conditions.

Table 5, were subjected to pathway enrichment analysis using the DIANA platform [92]. Among the identified pathways, the ECM-receptor interaction pathway showed the highest enrichment

significance (p-value = 0.000255), linked to two genes and two miRNAs (hsa-miR-514a-3p and hsa-miR-198). The glutamatergic synapse pathway was enriched (p-value = 0.033733), involving three genes and hsa-miR-490-5p. Figure 12 illustrates the association of these miRNAs with enriched pathways, with the ECM-receptor interaction pathway showing the highest prominence.

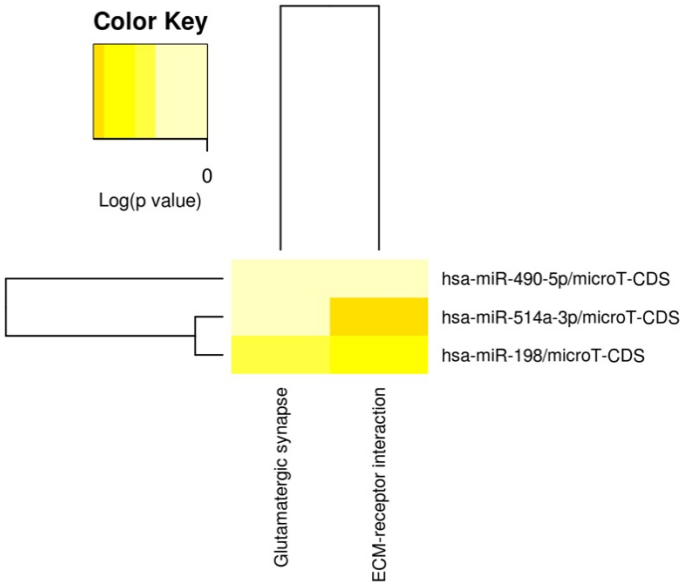


Figure 12. Heatmap of Enriched Pathways for Statistically Significant MicroRNAs Identified Using the DIANA Tool. Heatmap showing the association between significant microRNAs (hsa-miR-490-5p, hsa-miR-514a-3p, and hsa-miR-198) and their enriched pathways. Color intensity represents enrichment significance (log-transformed p-values), with darker shades indicating stronger significance. The ECM-receptor interaction pathway is notably enriched for hsa-miR-514a-3p and hsa-miR-198 (p-value = 0.000255), while the glutamatergic synapse pathway is linked to hsa-miR-490-5p (p-value = 0.033733).

2.3.2 AugPred Ablation study results

The results of the ablation study, showing as both the augmentation and the feature selection are crucial in obtaining stable and robust results, are summarized in Table 6

Method	Accuracy	False infection	False rejection
Baseline Approach	75%	2	1
AugPred - Logistic	82%	1	1
AugPred - KSVM	91%	0	1

Table 6. Performance comparison between the baseline logistic regression classifier (Baseline approach), simplified version of the pipeline with a logistic regression classifier (AugPred - Logistic), and the AugPred pipeline with KSVM classifier (AugPred-KSVM).

The pipeline with the KSVM classifier shows outstanding results compared to the baseline approach, which achieves a limited accuracy of 75%. Moreover, a clear improvement is observed with kernel-based classification compared to the logistic regression version.

2.3.3 AugPred Experiment 2 Results: Data scarce scenario simulation on E-MATB 8026 dataset results

Data scarce Cross-validation results

To evaluate the classification accuracy of the AugPred pipeline with a robust sample size, we conducted a cross-validation analysis. For each scenario (Healthy control vs LCa, Healthy control vs NTLD, LCa vs NTLD), 80 samples (40 from each class) were randomly selected and subjected to 4-fold cross-validation (k=4). This approach enabled a comprehensive assessment of the pipeline's performance under controlled conditions, ensuring a balanced dataset with adequate representation from both classes in each scenario.

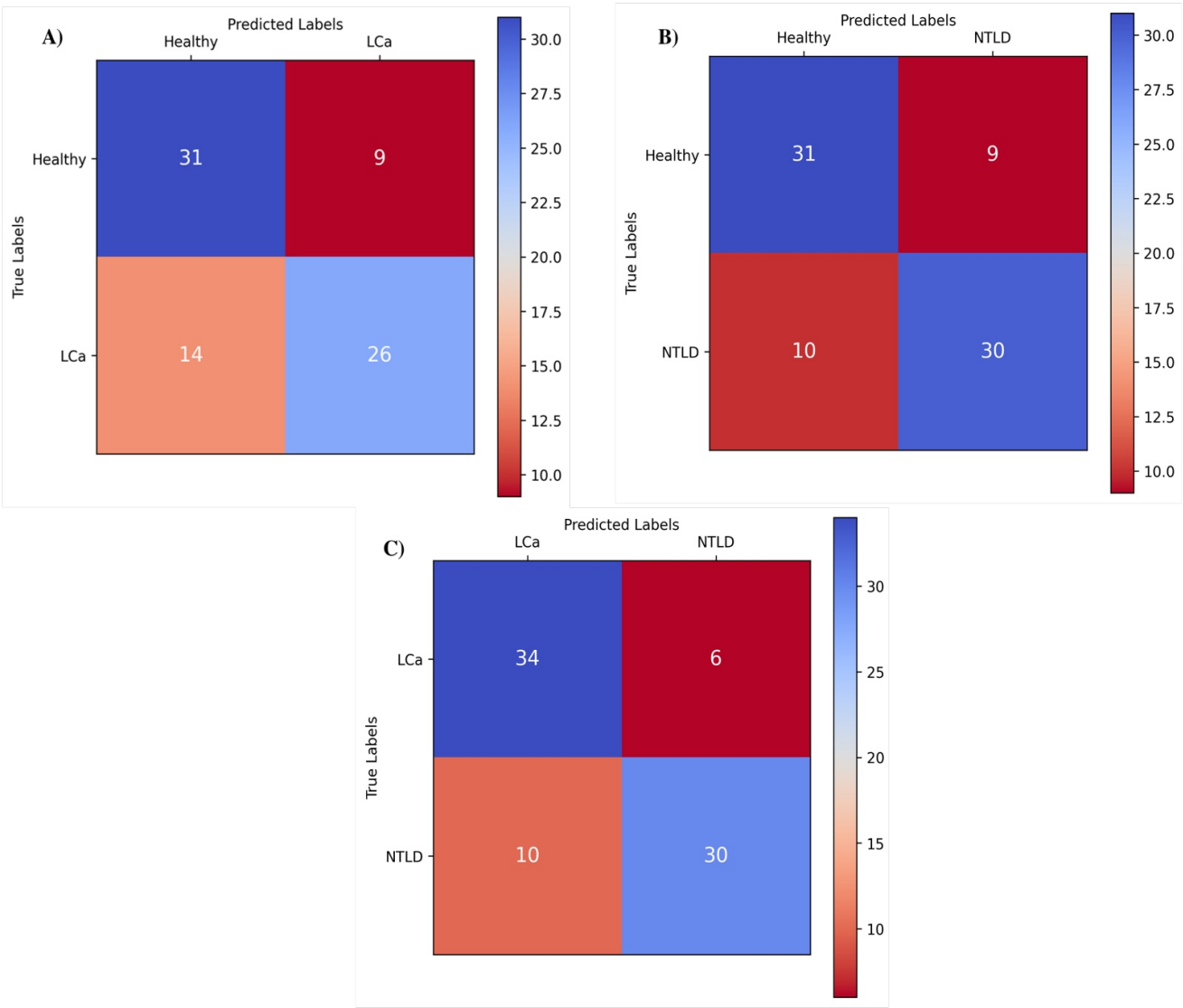


Figure 13. Cross-validation accuracy for each scenario. Confusion matrices illustrating the performance of the AugPred pipeline in three binary classification scenarios: (A) Healthy controls vs. Lung Cancer (LCa), (B) Healthy controls vs. non-tumor lung disease (NTLD), and (C) LCa vs. NTLD. Each scenario involved 80 samples (40 per class) subjected to 4-fold cross-validation with a 70%/30% training and testing split. The pipeline achieved an accuracy of 71% for Healthy vs. LCa, 76% for Healthy vs. NTLD, and 80% for LCa vs. NTLD. Color intensity and associated color-bar indicate the count of samples in each cell, with darker shades representing higher values.

Figure 13 shows the confusion matrices for the three classification scenarios tested with the AugPred pipeline. In the Healthy controls vs. LCa scenario, the pipeline achieved an accuracy of 71% (Figure 5 - A), correctly classifying most samples while misclassifying 14 LCa cases as healthy and 9 healthy controls as LCa. For Healthy controls vs. NTLD, the performance improved to 76% (Figure 5 - B), with balanced misclassifications between the groups, suggesting clearer separation in miRNA expression profiles. The best performance was observed in LCa vs. NTLD, with an accuracy of 80% (Figure 5 - C), where only 6 LCa and 10 NTLD samples were misclassified.

Generalization results

After the AugPred pipeline has been trained on a subset of 80 samples (40 per class), it is tested on all remaining samples in the dataset (testing samples per each class: 566 LCa, 553 NTLD, and 924 healthy control).

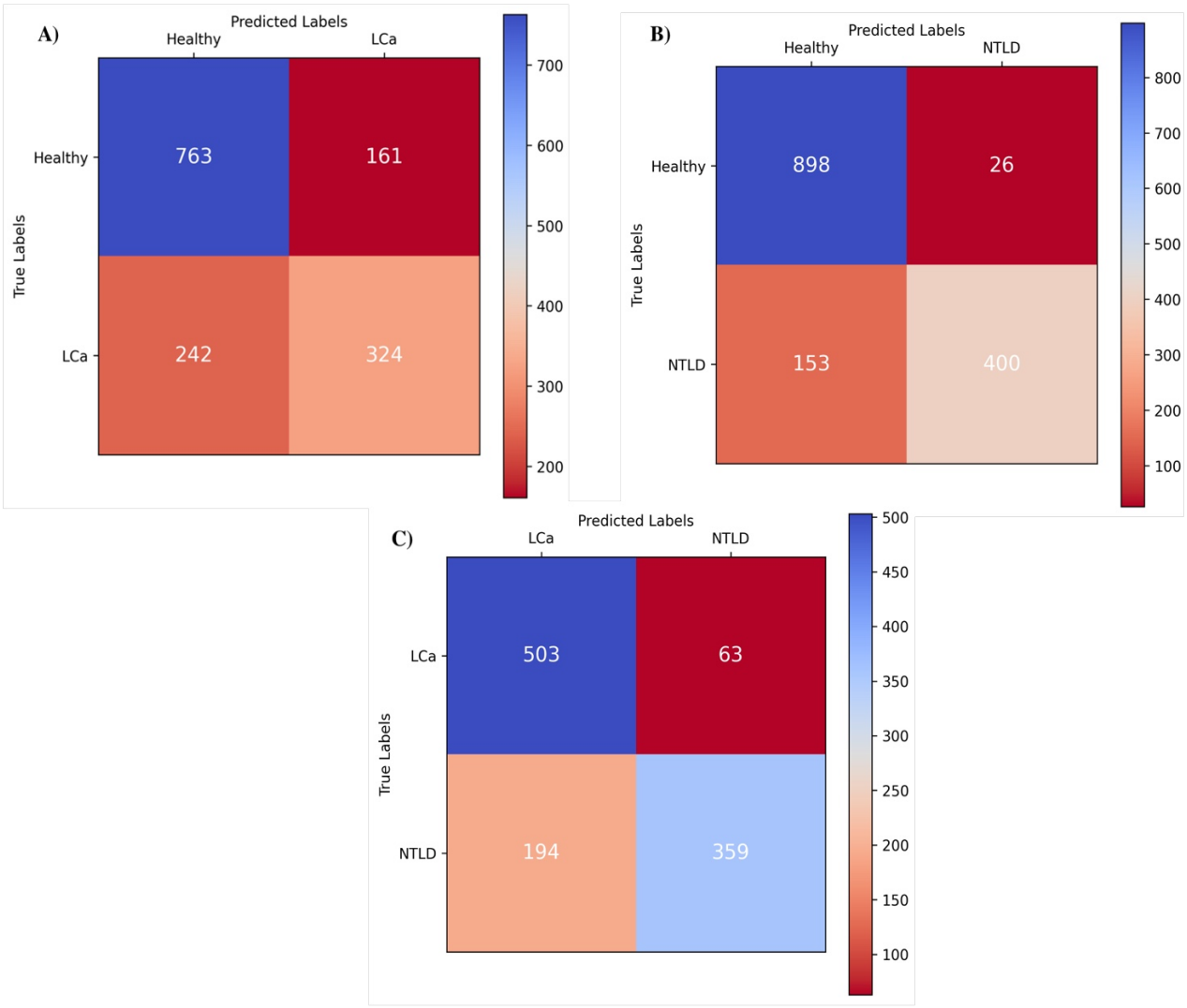


Figure 14. Generalization accuracy performance for each scenario. Confusion matrices showing the performance of the AugPred pipeline on the full dataset across three binary classification scenarios: (A) Healthy controls vs. Lung Cancer (LCa), (B) Healthy controls vs. Non-tumor Lung Disease (NTLD), and (C) LCa vs. NTLD. The model was trained on a subset of 80 samples (40 per class) and tested on all remaining samples (964 healthy controls, 606 LCa cases, and 593 NTLD cases). The pipeline achieved an accuracy of 73% for Healthy vs. LCa, 88% for Healthy vs. NTLD, and 77% for LCa vs. NTLD. Color intensity represents the count of samples in each cell, with darker shades indicating higher values.

Figure 14 presents the confusion matrices for the three classification scenarios. The results demonstrate stable and promising levels of accuracy across the scenarios. In the Healthy controls vs. LCa scenario, the pipeline achieved an accuracy of 73%, with most healthy controls and LCa cases correctly classified, despite some misclassifications on both sides. For the Healthy controls vs. NTLD scenario, the classifier achieved its best performance, with an accuracy of 88%, maintaining strong consistency with the cross-validation phase. The distinct separation of miRNA profiles between healthy controls and NTLD cases contributed to the minimal misclassifications observed in this scenario. In the LCa vs. NTLD scenario, the pipeline achieved an accuracy of 77%, effectively distinguishing between these two groups while managing some overlap in miRNA expression patterns. The results remain aligned with the cross-validation performance, further underscoring the model's robustness. When comparing the results between the cross-validation phase and the full dataset testing, consistent performance is evident across both evaluation methods. A slight but stable decrease in accuracy was observed for the Healthy controls vs. LCa (from 71% to 73%) and LCa vs. NTLD (from 80% to 77%) scenarios, while a significant improvement was noted for Healthy controls vs. NTLD (from 76% to 88%).

Extreme scarce data generalization results

We randomly selected 10 samples per class (approximately 1.6% of the entire dataset) from the LCa vs. NTLD scenario, trained the AugPred pipeline models, and performed classification. This process was repeated 100 times to ensure reliable results. At each iteration, we recorded the confusion matrix, calculated the accuracy, and computed the average accuracy across all iterations. The results show an average accuracy of 70.3% (Figure 15). The slight decrease in accuracy compared to the cross-validation and the previous study using 80 training samples highlights the robustness of the classifier and its ability to generalize, even when trained on a highly limited sample size.

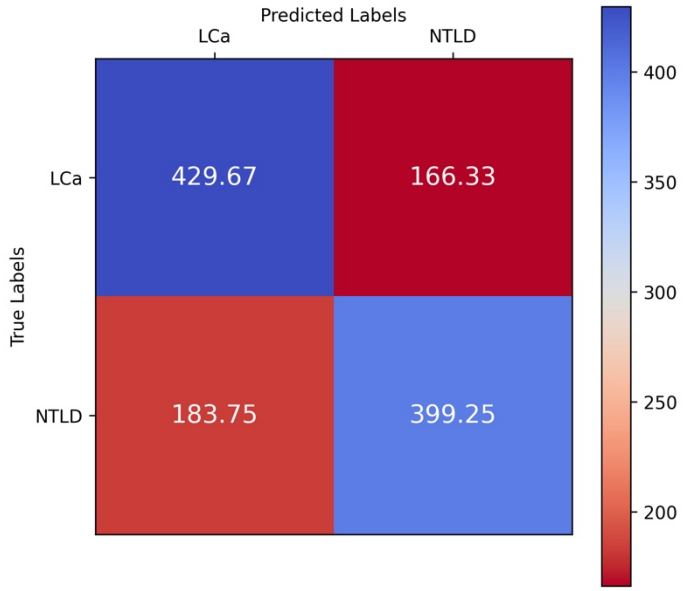


Figure 15. Average confusion matrix robustness analysis. Average confusion matrix obtained from 100 iterations of the AugPred pipeline, where the model was trained on 10 randomly selected samples per class (1.6% of the dataset) in the LCa vs. NTLD scenario. The average accuracy across iterations was 70.3%

2.4 Discussion

The work we have presented stems from the methodological foundation laid by the bootstrap analysis, which aimed to evaluate classification performance in omics datasets under varying data availability conditions. This initial exploration not only validated the effectiveness of integrating feature selection and synthetic data generation techniques but also revealed the limitations of classical models in terms of interpretability and robustness. The insights gained through this study paved the way for the design and implementation of the AugPred pipeline, which is the targeted solution employed for phenotype classification in the post-heart transplant setting.

Bootstrap Analysis

The results of the bootstrap analysis demonstrate the effectiveness of integrating feature selection and synthetic data generation in omics-based classification tasks. The baseline LASSO regression model exhibited relatively better classification accuracy across multiple scenarios, as shown in Figure 2 and Table II. However, this came at the critical cost of retaining a significantly larger number of features (Table I), which can reduce model interpretability and increase the risk of overfitting, particularly in high-dimensional datasets. The proposed L1-KSVM framework offered an optimal balanced trade-off between high accuracy and feature selection [93]. The use of data augmentation significantly enhances classification performance compared to its non-augmented counterpart, particularly in low sample conditions, where the model benefits from the additional synthetic training examples, improving its generalization ability. Moreover, the L1-KSVM framework achieved comparable accuracy to the baseline model while drastically reducing the number of selected features, making it a more interpretable and explainable solution. We decided to include in the framework the L1-KSVM architecture to evaluate a more expressive model capable of capturing non-linear patterns through kernel-based projections while maintaining feature selection through L1 regularization. This combination aimed to balance model complexity and interpretability. The use of synthetic data via Gaussian noise further tested whether data augmentation could improve robustness in low-sample settings [94]. We demonstrate that the proposed frameworks achieve consistent cross-validated performance on small datasets that emulate real-world scenarios. The accuracy remains stable when the trained classifier is applied to a larger test set, highlighting its robustness and generalizability. Reducing high dimensionality by selecting relevant features, which enables accurate classification of patient conditions, provides a significant advantage in understanding underlying biological insights and facilitates more precise statistical and bioinformatics analyses. This study advances the integration of synthetic data into omics-based analysis while improving decision interpretability through targeted feature selection. Future research will aim to validate the framework across multiple omics datasets and explore more sophisticated data augmentation techniques, such as generative deep learning models (like generative adversarial networks (GANs) and variational autoencoders (VAE)) [95] [96] to design more biologically meaningful synthetic samples. This approach will be crucial in overcoming the study's limitation, as Gaussian noise-based augmentation may not fully capture the complex biological variability present in real-world omics datasets.

The AugPred Pipeline

Building upon the findings of the bootstrap study, we developed and validated AugPred, a machine learning pipeline specifically designed to leverage microarray data for phenotype classification in post-heart transplant surveillance.

The pipeline, based on miRNA expression on EMBs, effectively distinguishes ACR from CMV infections, achieving 91% classification accuracy in this critical clinical context. In detail, AugPred identified three miRNAs as the most statistically significant in classifying EMB samples as CMV or ACR: hsa-mir-514a-3p, hsa-mir-198, and hsa-mir-490-5p. Hence, these miRNAs are the best candidates for future validation studies in a larger patient cohort. Furthermore, by integrating data augmentation techniques, L1-regularized logistic regression for feature selection, and classification through KSVM, AugPred significantly outperforms traditional approaches, demonstrating its potential to enhance diagnostic precision in high-dimensional omics datasets derived from heart tissue biopsies. The miRNAs showing the most significant differences in distribution between CMV and ACR groups (hsa-mir-514a-3p, hsa-mir-198, and hsa-mir-490-5p) have been analyzed exploiting the DIANA pathway enrichment analysis tool. The results highlighted that hsa-miR-514a-3p and hsa-miR-198 are linked to the ECM-receptor interaction pathway, which showed the highest enrichment significance. The extracellular matrix (ECM)-receptor interaction pathway plays a crucial role in mediating cell-ECM communication, which is essential for cellular adhesion, migration, and tissue remodeling [97][98]. This pathway is particularly relevant in the context of transplantation, where it contributes to graft remodeling and repair processes. Dysregulation of this pathway, such as overexpression of ECM components or their receptors, has been associated with pathological conditions including fibrosis and chronic allograft rejection [97]. Excessive ECM deposition and alterations in receptor signaling can lead to fibrotic tissue remodeling, a hallmark of chronic graft dysfunction and eventual failure [99].

In addition, hsa-miR-490-5p is reported to be involved in the Glutamatergic synapse pathway, which also displayed significant relevance in pathway enrichment analysis. The glutamatergic synapse pathway, although primarily associated with neurotransmission in the central nervous system, has been implicated in the regulation of immune responses during transplantation [100]. Activation of glutamate receptors on immune cells can influence cytokine production and inflammatory responses, potentially exacerbating graft rejection [21], [22]. These are potential key findings to better understand the molecular mechanisms underlying graft rejection and immune regulation in transplantation, providing a foundation for further studies on diagnostic biomarkers and therapeutic targets. Notably, some miRNAs that did not exceed the threshold of significance for signal intensity distribution difference between CMV and ACR groups but were selected from the data-driven feature selection approach, such as hsa-miR-1323, hsa-miR-30a-3p, hsa-miR-30a-5p, hsa-miR-10b-3p, and hsa-miR-10b-5p, are involved in cardiovascular pathophysiological mechanisms and pathways. For example, hsa-miR-1323, hsa-miR-30a-3p, and hsa-miR-30a-5p have been reported in other important studies in the heart transplant field. Indeed, Di Francesco et al. showed that these miRNAs are differentially expressed in ACR, antibody-mediated rejection, and mixed rejection [103]. Another interesting study, conducted by Zhong et al., demonstrated through qRT-PCR the role of multiple primary miRNAs precursor, including hsa-miR-30a and hsa-miR-10b, in playing an important role in regulation of vascular calcification pathways, which is a condition often encountered in heart transplant patients at risk of allograft rejection [104], [105]. All

these published results are a domain-specific demonstration that our pipeline is capable of extracting meaningful features from microarray data.

AugPred has been validated on the E-MATB-8026 dataset, demonstrating robust performance across multiple scenarios, highlighting its adaptability and generalization capabilities even under low data availability conditions. In detail, AugPred proved to be an efficient classification tool that can be applied to various patient populations, regardless of the pathological setting and the population size. The consistent accuracy across diverse classification challenges underscores AugPred's scalability and versatility in handling complex datasets and proves the powerful implication of generated synthetic samples and machine learning models to improve system performance and mitigate limitations posed by small sample sizes. Furthermore, the pipeline's independence from specific microarray kits or sequencing platforms enhances its flexibility and broad applicability across various experimental setups. All these characteristics make AugPred a strong pipeline, widely applicable in translational research and clinical practice.

Even though AugPred demonstrates optimal results with easy implementation and applicability, it has some limitations that warrant consideration. Its reliance on invasive EMBs restricts its routine use, and the pipeline depends heavily on high-quality datasets, making it sensitive to data biases. For this reason, further studies will focus on the relationship with the inclusion of multi-omics data and possible circulating transcripts. Future work on AugPred will focus on exploring the integration of molecular features with clinical data, to expand its applications and possibly reduce reliance on invasive procedures. Optimizing scalability and incorporating the pipeline into clinical workflows could further establish its role in personalized medicine. On the other hand, we need to enroll a new cohort of patients to validate the efficacy of AugPred-selected miRNAs as biological biomarkers in heart transplant monitoring. In conclusion, AugPred is a significant advancement in machine learning for transplant medicine, offering robust performance and biological insights. With future refinements and validations, it has the potential to become an essential tool for tackling complex clinical challenges.

3 Multilayer Networks Modeling of miRNA-mRNA Interactions

3.1 Aim of the study

Biological systems are inherently complex, characterized by a multitude of interacting molecular components whose relationships span different types, scales, and regulatory hierarchies. In recent years, complex network theory has emerged as a powerful paradigm to model such systems, providing tools to investigate their structural, functional, and regulatory organization [106,107]. In systems biology, complex networks have been extensively applied to represent and interrogate diverse molecular interactions. In this framework, molecules such as genes, RNAs, and proteins are represented as nodes, while their interactions whether physical, functional, or regulatory are described as edges. This abstraction enables the analysis of network topology, identification of key regulatory hubs, and exploration of emergent behaviors across biological contexts [108]. Biological regulation is inherently complex, involving multilayered and interdependent relationships among molecular entities. In this context, microRNAs (miRNAs) have emerged as pivotal post-transcriptional regulators. These small, non-coding RNA molecules bind to target messenger RNAs (mRNAs), modulating their stability and translation. Through this mechanism, miRNAs orchestrate essential cellular processes, including immune activation, cell cycle control, and apoptosis. Their role becomes particularly critical in complex clinical scenarios such as heart transplantation, where regulatory imbalances may contribute to adverse outcomes [109]. Among post-transplant complications, two conditions are especially relevant: acute cellular rejection (ACR) and cytomegalovirus (CMV) infection. ACR, the most common form of rejection within the first year after transplant, arises from immune-mediated responses due to mismatches between donor and recipient major histocompatibility complexes (MHCs) [110]. CMV infection, on the other hand, represents the most clinically significant viral complication in transplant recipients [111]. Although ACR and CMV infection are biologically distinct, one immune-mediated and the other viral, they often exhibit overlapping histopathological features, complicating differential diagnosis and treatment. In this context, miRNAs have gained attention as promising molecular markers. Understanding how miRNAs and their mRNA targets behave under these divergent but intersecting conditions is crucial to uncovering their regulatory mechanisms [112] [113]. Yet, most transcriptomic studies in transplantation rely primarily on differential expression analysis, which, while informative, treats molecules as independent units [114]. This approach overlooks the interconnected nature of gene regulation, where individual miRNAs may regulate multiple mRNAs, and each mRNA may be targeted by several miRNAs, forming complex regulatory networks. To investigate whether regulatory influence shifts across different post-transplant conditions, we constructed multilayer networks integrating miRNA and mRNA expression data from endomyocardial biopsies of 16 cardiac transplant recipients. Intra-layer connections were defined based on ontology-driven functional similarity among miRNAs and mRNAs, while inter-layer edges captured high-confidence predicted interactions between miRNAs and their mRNA targets. We applied an ensemble PageRank strategy to robustly estimate the centrality of each node, capturing its structural and regulatory importance within the network. By comparing centrality profiles across control, infection, and rejection phenotypes, and performing pathway enrichment analyses on differentially central molecules, we aimed to uncover phenotype-specific signatures of post-transcriptional regulation that may not be detectable through expression analysis alone. This work offers both methodological and biological contributions, introducing a scalable, data-driven

framework for integrative transcriptomic analysis and providing novel insights into the molecular mechanisms distinguishing CMV infection from rejection. The proposed approach is also generalizable to other disease settings and adaptable to longitudinal designs for studying dynamic regulatory changes over time.

3.2 Dataset and Preprocessing

For this study, 16 patients who underwent cardiac transplantation between January 2018 and December 2020 were selected. We excluded pediatric patients, and adult patients that tested positive for coinfection. All participants were provided with complete information about the study, and all procedures performed were in accordance with the ethics standards of the institutional and/or national research committee and with the 2013 Helsinki Declaration and its later amendments or comparable ethical standards. Endomyocardial biopsies (EMBs) of these patients were collected. EMBs were carefully reviewed by two expert pathologists, who reached an agreement on the grading. EMBs were classified according to the ISHLT 2013 updated international classification criteria for antibody-mediated rejection [115] and the ISHLT 2005 working formulation for acute cellular rejection [116] and divided into 3 phenotypic classes:

- **5 Controls:** patients with negative EMB for rejection
- **6 Infections:** patients with cytomegalovirus infection (CMV).
- **5 Rejections:** patients experiencing acute cellular rejection (ACR) of the allograft.

Patients with biopsies showing inflammatory infiltrates are classified as acute cellular rejection (ACR) (Rejections), while patients with biopsies indicating acute rejection and microbiological evidence of CMV viremia exceeding 1,000 copies/mL are classified as CMV infection (Infections). According to the 2005 guidelines of the International Society for Heart and Lung Transplantation (ISHLT) [116], all rejection cases in the dataset were classified as 3A-2R. Among the infection cases, four were classified as 1A-1R, one as 3A-2R, and one as 3B-2R. Control samples include patients who were negative at EMB for rejection. Total RNA was extracted using RecoverAll Total Nucleic Acid Isolation Kit (Life Technologies, USA) according to the manufacturer's instructions. High-density Clariom S Affymetrix GeneChip arrays (Thermofisher Scientific, USA) were used to measure mRNA expression profiles. On the same total RNA extracted, microRNA expression profiles were measured using the Clariom-S miRNA 4.0 chips and FlashTag^{TM} Biotin HSR RNA Labeling Kit (Thermofisher Scientific, USA) following the manufacturer's instructions. Measured expression profiles of mRNAs and microRNAs were preprocessed by filtering out missing values and entries with zero signal intensity. Additionally, to reduce the influence of background noise and low-confidence measurements, features with expression values below the 25th percentile across samples were removed. MicroRNAs without the 'hsa-' prefix were discarded to retain only human-specific microRNAs relevant to the study. After preprocessing, a total of 2,561 microRNAs and 3,688 mRNAs per sample were retained for downstream analysis

3.3 Methodology

The entire workflow of the developed multilayer network framework can be visualized in Figure 16 which outlines the three main stages of the analyses: (A) data preprocessing, (B) multilayer network construction and ensemble PageRank computation, and (C) PageRank analyses and biological interpretation. The multilayer network model and the ensemble PageRank computation were implemented using R programming language. The multilayer network model and the ensemble PageRank computation were implemented using the R programming language, ensuring reproducibility and flexibility in graph manipulation and centrality analysis. The pathway enrichment analysis (PEA) was performed using the EnrichR R package, which provides access to multiple pathway and gene set libraries, including KEGG. All statistical analyses, data integration procedures, and the microRNA–pathway association study were conducted using R and Python, leveraging libraries such as pandas for data handling, statsmodels for statistical testing, and ggplot, matplotlib and seaborn for visualization.

3.3.1 Multilayer Network

Multilayer Network Modelling

By representing molecules as nodes and multiple types of interactions as edges, complex network modelling allows for the investigation of regulatory, functional, and structural relationships within cellular processes. To address the complexity of regulatory interactions between microRNAs and mRNAs across three categories (control, infection, and rejection), we propose a multilayer network framework to analyze and integrate these relationships. Multilayer networks enable the simultaneous representation of different types of biological interactions within a unified and structured model [117,118]. Intra-layer connections (or Intra-links) describe interactions within the same molecular class (e.g., microRNA–microRNA or mRNA–mRNA associations), while inter-layer connections (or Inter-links) represent regulatory links between microRNAs and their target mRNAs. A representation of the constructed multilayer microRNA–mRNA network is shown in Figure 18. In our proposed methodology the multilayer network Intra-links between nodes of the same molecular type were constructed following GO-based functional and semantic similarity, to capture biological roles relationships. Specifically, microRNA–microRNA connections were defined using MirGofos [119], an unsupervised method that infers functional similarity between microRNAs based on the semantic similarity of their gene ontology (GO)-annotated target genes. mRNA–mRNA associations were computed using GOGO [120], a graph-based algorithm that quantifies the semantic similarity between gene products by exploiting the topological structure of the GO graph. Together, these methods ensure that intra-layer connections reflect weighted shared biological functions.

Inter-layer Connections and Biological Assumptions.

Inter-layer connections (Inter-links) represent the regulatory interactions between microRNAs and their target mRNAs. Most of the microRNAs – mRNAs kinetic interaction modelling approaches rely on correlation-based approaches or mathematical modelling of co-expression patterns, eventually including target prediction databases to infer regulatory relationships.
In this study, we adopted a probabilistic approach to exploit the static expression profiles of each

sample availability. Therefore, our objective is to estimate the potential for interaction between miRNAs and their target genes using expression data collected at a single time point (t = 0) [121,122]. Although miRNA-mediated regulation is inherently dynamic and context-dependent**,** we assume that an increase in the concentration of a given miRNA, combined with a high predicted affinity for a specific mRNA target**,** is indicative of a potential negative regulatory interaction via the RISC complex**.** To mitigate the risk of false-positive predictions arising from the absence of temporal data, we considered only interactions with a Target Prediction Score (Sc) ≥ 80, as provided by miRDB, a threshold identifying highly reliable predictions based on high-throughput experimental data [123]. Nevertheless, we acknowledge that the likelihood of miRNA–mRNA interaction is also influenced by additional factors**,** such as competition with other targets (ceRNA effect), the availability of RISC complex proteins, and cell-type-specific contexts [124] [125]. To quantify the relative probability of interaction of each pair of microRNA-mRNA, we computed an Interaction Score (IS) based on the measured concentrations of miRNAs and their target mRNAs, together with the Sc value (2. This score was then normalized according to the distribution of predicted interactions for each microRNA Eq. 3 yielding the Interaction Probability (PI), which represents the relative likelihood of regulation for each miRNA-mRNA pair.

$$IS_{i,j} = [\,]microRNA_i \times [\,]target_j \times Sc_{i,j} \tag{2}$$

$$PI_{i,j} = \frac{IS_{i,j}}{\sum_{t=1}^{n} IS_{i,j}} \tag{3}$$

3.3.2 Ensemble PageRank Approach

PageRank is a commonly used centrality measure initially developed to rank webpages in the context of the World Wide Web [65]. It quantifies the relative importance of nodes in a directed graph by considering both the number and the quality of connections they receive. In the context of biological networks, PageRank can be used to infer the regulatory influence or topological relevance of a given molecule (e.g., a miRNA or mRNA) within a larger regulatory architecture. The intuition behind PageRank is that a node is considered important if it is linked to by other important nodes. This recursive formulation allows PageRank to capture indirect regulatory influence, making it particularly suitable for identifying key regulators in complex and interconnected biological systems. Formally, PageRank is computed as the steady-state distribution of a random walk process with damping, where the probability of transitioning from one node to another is determined by the network's connectivity structure. The iterative process converges to a ranking vector, where each element reflects the long-term visitation probability of the corresponding node. To assess node centrality within the multilayer microRNA–mRNA regulatory network, we implemented an ensemble PageRank strategy designed to produce a robust and stable

estimation of node importance, while mitigating the computational burden of computing PageRank across the entire multilayer network in a single step. Instead of evaluating the full network at once, we employed an iterative ensemble procedure that progressively stabilizes the PageRank distribution through multiple randomized subnetwork constructions. Specifically:

- **Initialization:** PageRank values were initialized to zero for all microRNAs and mRNAs. Two additional vectors were used to track convergence across iterations.

- **Sampling:** At each iteration, a subset of $\theta = 512$ microRNA nodes (approximately 20% of the total) was randomly sampled from the network.

- **Target Retrieval:** For each sampled microRNA, all associated target mRNAs were retrieved based on non-zero values in a sample-specific transition probability matrix.

- **Subnetwork Construction:** multilayer subnetworks were created, including related intralayer and interlayer connections.

- **PageRank Computation:** PageRank scores were computed for the nodes in the sampled subnetwork and aggregated into the global PageRank vector.

- **Convergence Monitoring:** The mean absolute percentage difference (or mean absolute percentage error, MAPE) of PageRank was tracked as convergence metric.

- **Stopping Criterion:** The process continued until either the maximum number of iterations ($M = 2000$) was reached, or convergence was detected, defined as 25 consecutive iterations where the absolute relative percentage change in PageRank fell below 10^{-6}.

Figure 17 shows the convergence of the ensemble PageRank process, based on the absolute relative percentage change in PageRank values for microRNAs (microRNA) and mRNAs.

3.3.2 PageRank and Bioinformatic Analyses

PageRank Distribution Analyses

PageRank value distributions were computed for both microRNAs and mRNAs across all phenotypic classes (control, infection, and rejection), following the convergence of the ensemble PageRank procedure. For each node, the mean PageRank score was calculated and normalized between 0 and 1, within each phenotype group. The resulting frequency distributions were analyzed to characterize the overall topological organization of the multilayer regulatory network and to assess potential differences in centrality patterns across phenotypes. To identify nodes exhibiting statistically significant differences in PageRank distributions across phenotypes, we performed pairwise comparisons of PageRank values for each node. Specifically, the Wilcoxon rank-sum test (Mann–Whitney U test) was applied to assess whether the PageRank distributions differed

significantly between phenotypic groups (control vs. infection, control vs. rejection, and infection vs. rejection). This non-parametric test was chosen due to its robustness in detecting distributional differences without assuming normality.

Pathways Enrichment Analyses

Pathway enrichment analysis (PEA) was performed on the set of nodes that showed statistically significant differences in PageRank centrality across the three phenotypic classes (control, infection, and rejection). The analysis was conducted using the EnrichR R package and the KEGG database (Kyoto Encyclopedia of Genes and Genomes) [126], to identify biological pathways potentially associated with phenotype-specific regulatory patterns. All pathways with a p-value < 0.05 were retained and annotated according to their original KEGG classification. Top 20 enriched pathways were selected based on adjusted p-value statistical significance (Benjamini–Hochberg correction) to focus on the most statically significant pathways and facilitate comparative analysis and visualization, while avoiding redundancy among highly similar functional annotations.

MicroRNAs – Pathways Association Study

MicroRNAs that showed statistically significant differences in PageRank value distributions across phenotypes were mapped to the top 20 enriched pathways, based on their predicted target genes. To enhance the robustness of the pathway–microRNA associations, a median-based divergence filter was applied. For each microRNA, the median PageRank was computed per phenotype, and pairwise divergence scores were defined as the absolute difference between these medians across the three phenotype comparisons. MicroRNAs with divergence scores below the overall median were excluded, retaining only those with higher phenotype-specific centrality variation. The filtered microRNAs were then grouped by associated pathways and stratified by phenotypic class (control, infection, and rejection). For each pathway, the mean PageRank distribution of the associated microRNAs was computed separately for each phenotype. The resulting distributions were then compared and statistically evaluated using the Wilcoxon–Mann–Whitney test, to identify pathways showing significant differences in microRNA regulatory activity across conditions.

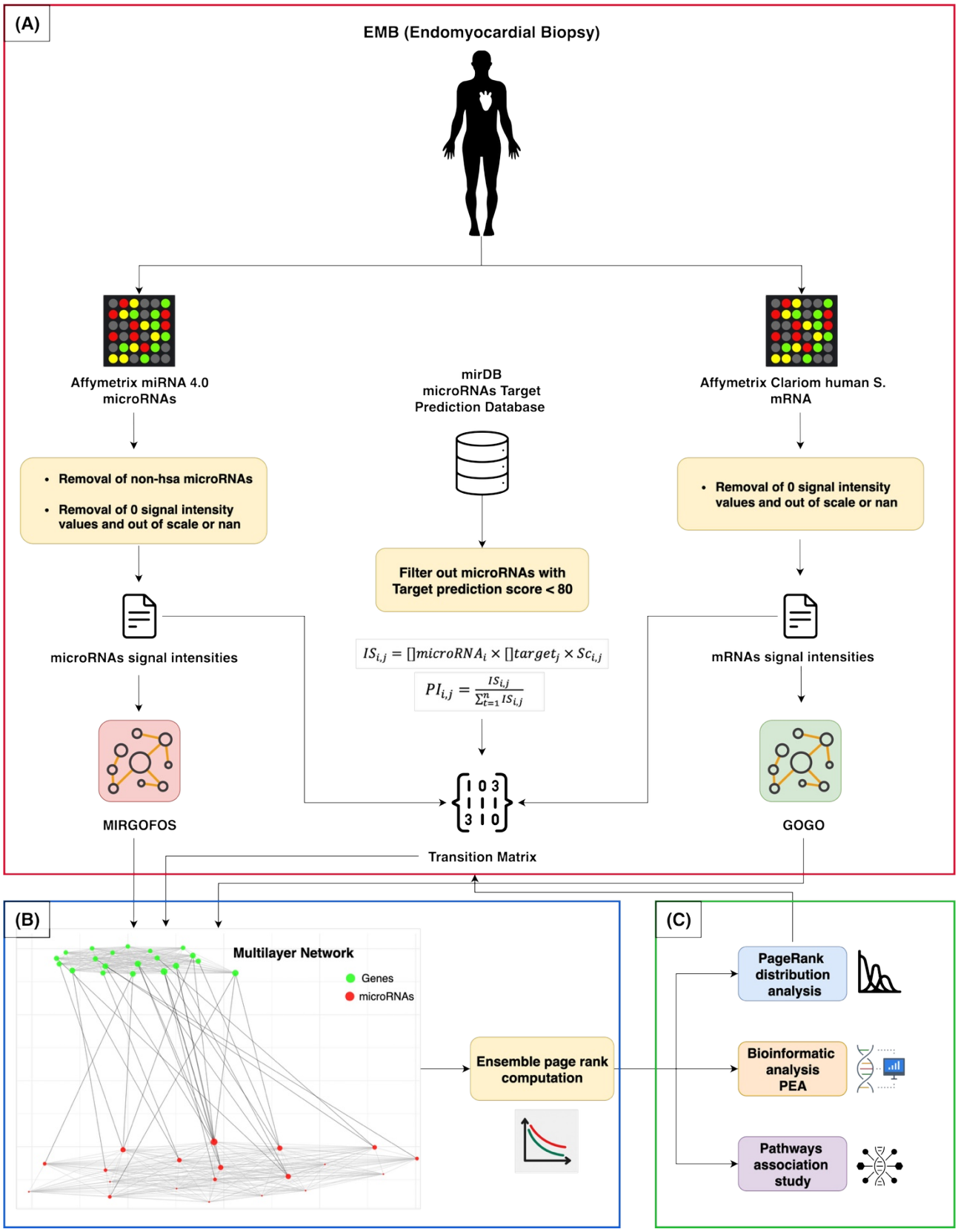


Figure 16 Multilayer Network-Based Framework Workflow. (A) Data preprocessing and network construction: microRNA and mRNA expression profiles are preprocessed, and multilayer network is constructed (B) PageRank computation: node centrality is assessed through an ensemble PageRank approach to ensure robust ranking. (C) PageRank analyses and Biological Interpretation: comparison of PageRank distributions across phenotypes, pathway enrichment analyses of the microRNAs or mRNAs with statistical relevant difference among phenotypes, and pathway–phenotype association studies

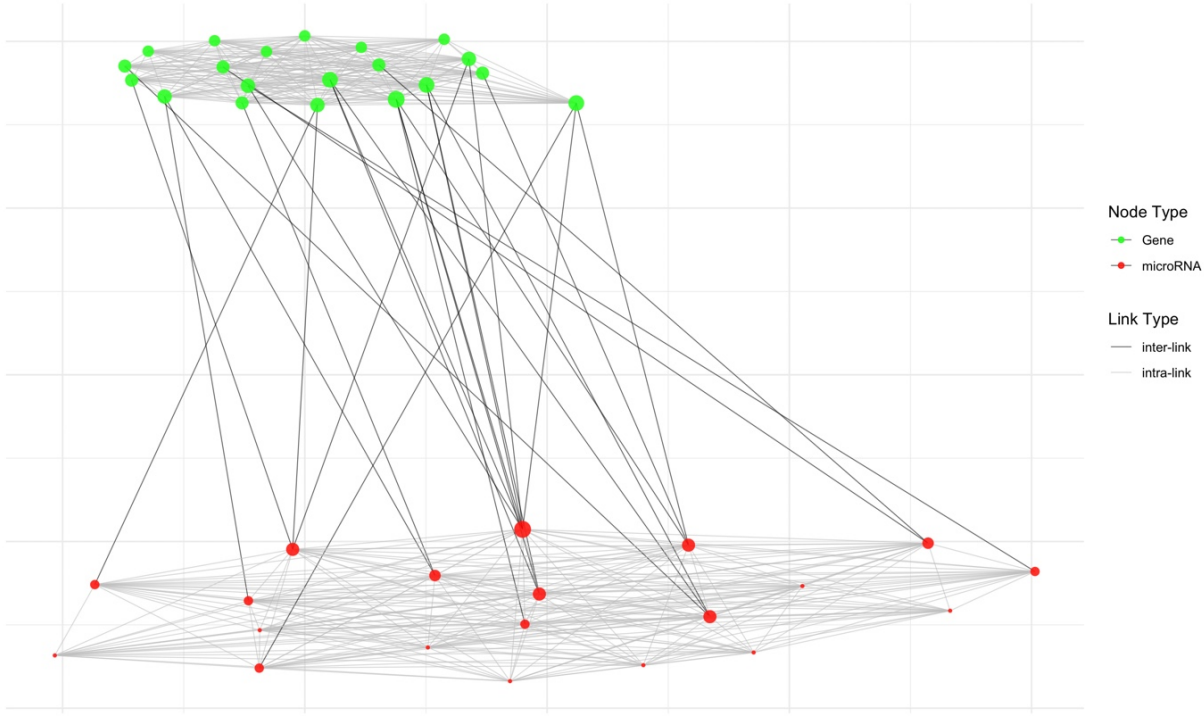


Figure 18. Multilayer complex network mRNA – microRNAs Graphical representation of the multilayer network integrating mRNAs (green nodes) and microRNAs (red nodes). Since mRNAs are usually associated to related transcribed gene, for simplicity we indicate them also as gene or target gene. its Intralayer connections represent interactions within the same molecular type (intra-link in grey), while interlayer connections capture microRNA–mRNA regulatory relationships (inter-link in black).

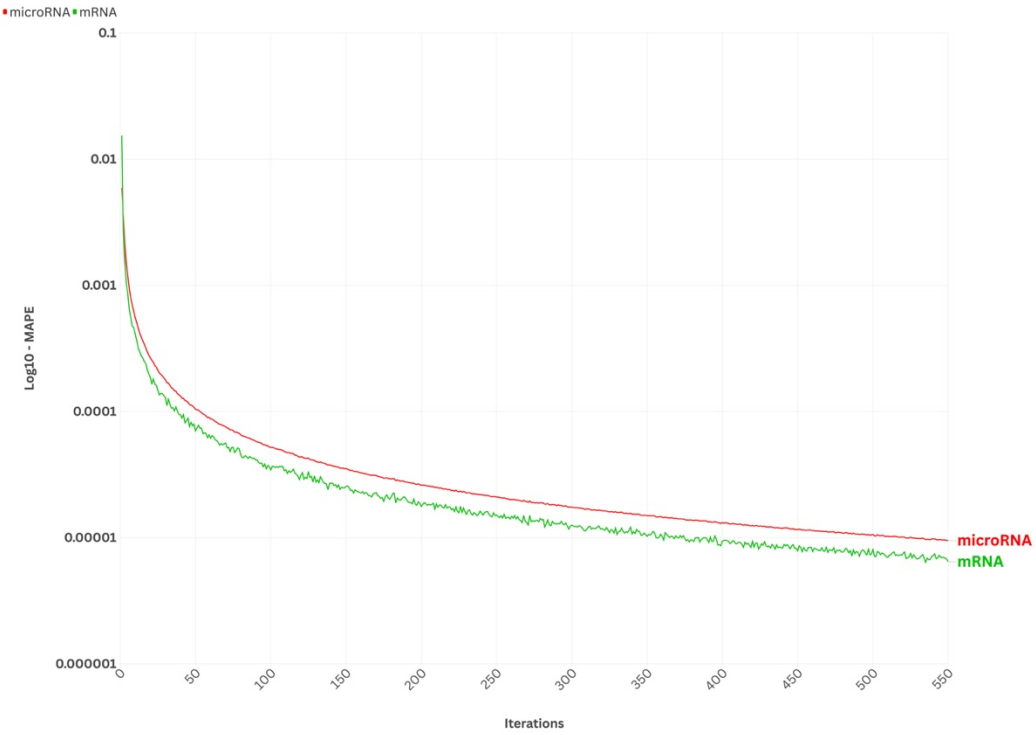


Figure 17. Convergence analysis of the ensemble PageRank computation for microRNAs and genes.
Computation of log10-scaled mean absolute percentage error (MAPE) for microRNA (red) and mRNAs (green) nodes during the ensemble PageRank estimation process. MAPE quantifies the difference between successive PageRank vectors at each iteration, demonstrating progressive convergence toward stable centrality values. Both microRNA and gene scores show consistent decay, with gene nodes converging slightly faster than microRNAs.

3.4 Results

3.4.1 PageRank Distribution Analysis

PageRank Distribution per Node Type and Phenotype

Mean PageRank values for each node were aggregated per phenotypic class (control, infection, rejection) and separated by molecular layer (mRNAs and microRNAs). Figure 19 displays the log-scaled normalized frequency distributions of log value of the PageRank centrality values across all conditions. A clear difference in the PageRank distributions was observed. mRNAs tend to have a broader distribution on higher level of PageRank, while microRNAs exhibited a left-skewed distribution a bit smaller PageRank values. When comparing distributions across phenotypic classes, no major shifts in centrality distribution were detected. The curves for control, infection, and rejection groups remained highly overlapping within each molecular class, both for mRNAs and microRNAs. This indicates a preservation of the topological structure of the Multilayer Network across the different samples, with similar centrality profile distribution across phenotypes.

Pairwise Comparison Across Phenotypes

The pairwise comparison analysis identified a total of 157 microRNAs with statistically significant differences (p-value < 0.05) in PageRank distributions across at least one phenotype comparison. In contrast, no mRNAs (transcribed genes) showed statistically significant centrality differences, with p-values consistently above or on the limit with the significance threshold. The results are summarized in the heatmaps shown in Figure 20, where color intensity represents the normalized PageRank values (range 0–1) for each microRNA across all samples. The heatmap highlights distinct patterns of centrality variation across phenotypes, with several microRNAs exhibiting visible changes in intensity between groups. Clustering analysis revealed partial grouping of samples according to phenotypic class, particularly for infection and rejection conditions, suggesting a structured variation in microRNA centrality profiles.

3.4.2 Pathways Enrichment Analysis (PEA)

From the 157 microRNAs with statistically different PageRank value among the phenotypic classes, a total of 195 pathways were identified as significantly enriched, with a p-value threshold of < 0.05. Multiple enriched pathways were associated to signaling pathways (24), cellular functions regulation (24) and both viral and bacterial infections (22). A relevant number of pathways (30) also mapped to the regulation of cardiovascular, circulatory, and endocrine system. Pathways related to cancer were also highly represented (24), that reflects cell proliferation pathways. Metabolic pathways accounted for 23 enriched terms, encompassing glycan biosynthesis, amino acid metabolism, and lipid-related processes. The presence of these pathways suggests a metabolic reprogramming in response to stress, immune activation, or pathogen burden, which is often observed in both infectious and inflammatory conditions. Sixteen enriched pathways were associated with the nervous system and neurodegeneration. These include processes such as axon

guidance and synaptic signaling, as well as disease-related pathways. Remaining pathways included key mechanisms of antigen presentation, cytokine signaling, and immune cell activation. Their enrichment supports the involvement of classical immune effectors. Figure 21 A) displays the top 20 significantly enriched pathways, ranked by statistical significance. Among these, the most enriched terms include Endocytosis and Axon guidance, followed by several intracellular signaling pathways such as Hippo signaling, MAPK signaling, and PI3K-Akt signaling. Infection-related pathways such as Human papillomavirus infection, Herpes simplex virus 1 infection, and Staphylococcus aureus infection are also prominently represented, together with canonical cancer pathways (Pathways in cancer, Proteasome, p53 signaling) and neurodegenerative disease-related terms. The upset plot in Figure 21 B) displays the intersection of genes among the top 20 enriched KEGG pathways. Notably, the Herpes simplex virus 1 (HSV-1) infection pathway stands out by containing the highest number of genes without intersection with other pathways (335). This suggests a strong, specific enrichment of this viral infection signature, potentially pointing to a distinct set of regulatory targets or immune-related responses influenced by the identified microRNAs. In contrast, several signaling pathways including the mTOR signaling pathway, MAPK signaling, Hippo signaling, and Ubiquitin-mediated proteolysis display intersecting gene sets, suggesting that these pathways might represent more general cellular programs affected during either infection, rejection, or stress responses post-transplant. The overlap between pathways in cancer, focal adhesion, and cellular senescence further indicates the involvement of pathways that regulate cell proliferation, apoptosis, and tissue remodeling, all relevant in the context of graft adaptation or injury.

3.4.3 MicroRNAs – Pathways association

To assess the potential regulatory influence of microRNAs on the most significantly enriched pathways, we mapped the 79 microRNA retained after median-based divergence filtering onto the pathways via their predicted target genes. These microRNAs demonstrated robust centrality profiles and were collectively associated with all the top enriched KEGG pathways. For each pathway, we then computed the mean PageRank value of the associated microRNAs within each phenotypic network (control, infection, and rejection).

As shown in Table 7, a general trend emerged in which the rejection phenotype exhibited lower average PageRank values for the majority of pathways, suggesting a reduced central role of these microRNAs in the regulatory network during acute rejection. Pathways such as Endocytosis, MAPK signaling, Pathways in cancer, and Protein processing in endoplasmic reticulum, which were also among the most significantly enriched and highly connected in terms of gene overlap, showed pronounced drops in mean PageRank in the rejection group. A similar pattern was observed for immune- and infection-related pathways like Herpes simplex virus 1 infection, Ubiquitin mediated proteolysis, and mTOR signaling, indicating that the microRNA-mediated regulation of these processes may be specifically diminished in the context of rejection.

Figure 19. Log–Log Distribution of Mean PageRank Values Across Phenotypes. The plot shows the normalized frequency distributions of mean PageRank values for mRNAs and microRNAs, stratified by phenotypic class (control, infection, and rejection). Both axes are log-scaled: the x-axis is log-transformed to emphasize the presence of long-tailed distributions in centrality values, while the y-axis is log-transformed to enhance visibility of variations in low-frequency regions.

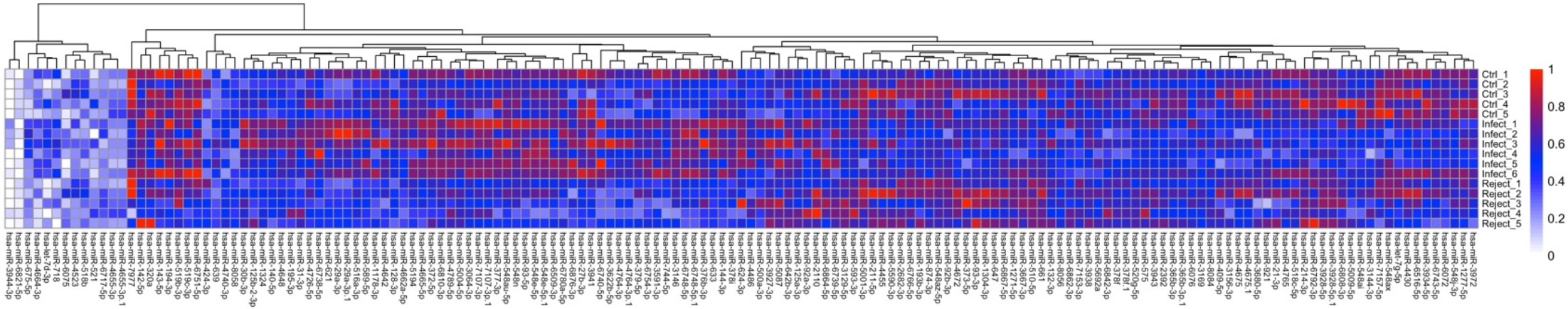


Figure 20. Normalized PageRank Values for Statistically Significant microRNAs. Heatmap displays the normalized PageRank values (range 0–1) for the 157 microRNAs that showed statistically significant differences (p-value < 0.05) in at least one phenotype comparison (control, infection, rejection). Each column corresponds to a sample, and each row to a microRNA. Color intensity represents the relative PageRank centrality, with red indicating higher values and blue indicating lower values. Hierarchical clustering on the columns highlights patterns of centrality variation across phenotype.

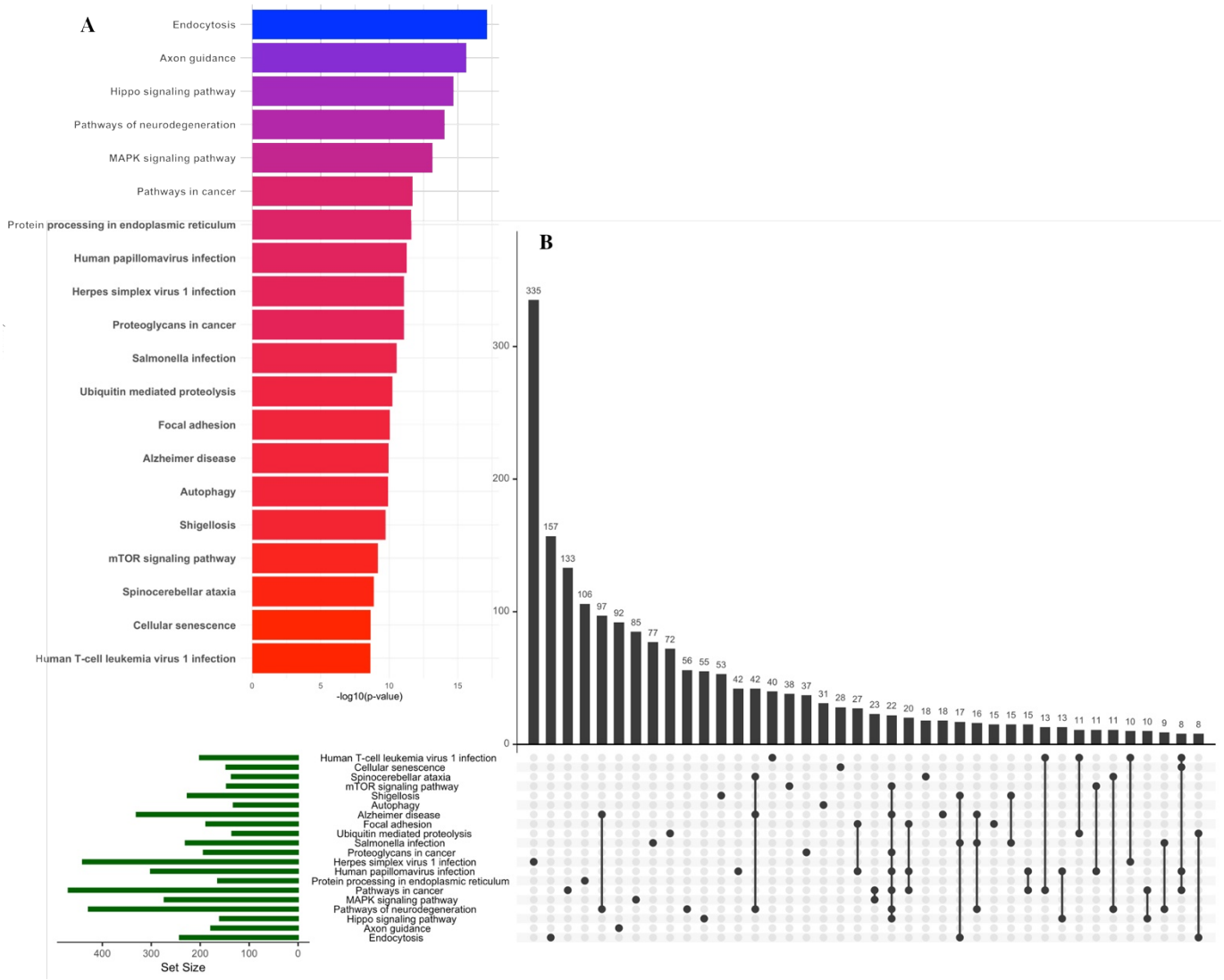


Figure 21. Top 20 Enriched pathways, Barplot and UpSet plot. (A) Barplot showing the 20 most significantly enriched pathways ranked by −log10(p-value). (B) UpSet plot showing gene set intersections across the top 20 enriched KEGG pathways. Horizontal bars (left) indicate the total number of genes annotated in each individual pathway. Vertical bars (top) represent the number of genes shared between pathway combinations indicated below, where black dots and connecting lines denote the specific intersections. The plot highlights both overlapping and unique gene content among the top 20 enriched terms, illustrating the functional connectivity between signaling, infection-related, and regulatory pathways.

Pathway	p-value control vs infection	p-value control vs rejection	p-value infection vs rejection	Mean PageRank Control	Mean PageRank Infection	Mean PageRank Rejection
Endocytosis	0.1045	2.94e-05	9.28e-08	0.6273	0.659	0.4976
Axon guidance	0.425	4.03e-05	1.56e-06	0.6363	0.6476	0.5075
Hippo signaling pathway	0.0988	8.42e-05	2.51e-07	0.6321	0.6573	0.4982
Pathways of neurodegeneration	0.1014	2.48e-05	5.81e-08	0.6225	0.6573	0.4982
MAPK signaling pathway	0.1367	4.31e-05	1.37e-07	0.6225	0.6504	0.4982
Pathways in cancer	0.0769	2.66e-05	3.42e-08	0.6182	0.6573	0.497
Protein processing in endoplasmic reticulum	0.1325	2.26e-05	8e-08	0.6322	0.6539	0.4871
Human papillomavirus infection	0.079	5.57e-05	7.43e-08	0.6182	0.6573	0.497
Herpes simplex virus 1 infection	0.1992	4.37e-05	3.61e-07	0.6321	0.6504	0.4982
Proteoglycans in cancer	0.0619	3.39e-05	4.19e-08	0.6182	0.6573	0.497
Salmonella infection	0.0964	3.05e-05	6.15e-08	0.6273	0.659	0.4976
Ubiquitin mediated proteolysis	0.2726	0.000103	2.29e-06	0.6322	0.6476	0.497
Focal adhesion	0.1253	0.000227	9.95e-07	0.6124	0.6504	0.4982
Alzheimer disease	0.1777	3.86e-05	2.43e-07	0.6321	0.6504	0.5075
Autophagy	0.0923	8.75e-05	2.91e-07	0.6321	0.6715	0.5075
Shigellosis	0.1016	3.72e-05	9.86e-08	0.6225	0.6573	0.497
mTOR signaling pathway	0.3002	5.48e-05	1.78e-06	0.6322	0.6476	0.4982
Spinocerebellar ataxia	0.2504	0.000207	5.09e-06	0.6343	0.6435	0.5029
Cellular senescence	0.1015	3.19e-05	6.78e-08	0.6203	0.6539	0.4976
Human T-cell leukemia virus 1 infection	0.1879	5.68e-05	3.53e-07	0.6273	0.649	0.5029

Table 7. The table summarizes p-values for pairwise comparisons of mean PageRank values of microRNAs associated with each pathway, across control, infection, and rejection groups. For each pathway, mean PageRank values in each phenotype are reported

3.5 Discussion

Our results revealed that the overall topological centrality profile of the microRNA–mRNA regulatory multilayer network remains consistent across phenotypes. This observation reflects a biologically meaningful organization, where transcribed mRNAs retain a higher centrality role, likely due to the convergent regulatory action of multiple microRNAs targeting the same gene [127]. A pairwise comparison of PageRank centrality revealed a clear divergence between microRNAs and mRNAs. While the centrality profiles of mRNAs remained consistent across phenotypes, showing no statistically significant differences, 157 microRNAs exhibited significant shifts in their centrality values ($p < 0.05$) in at least one phenotype comparison. The heatmap in Figure 20 highlights how the PageRank centrality of the selected microRNAs varies markedly across control, infection, and rejection phenotypes, revealing distinct expression and regulatory patterns. This differential behavior highlights microRNAs as key phenotype-specific regulators within the multilayer network, suggesting that they dynamically reshape their regulatory influence in response to phenotypic changes [128] [129]. Pathway enrichment analysis of the 157 microRNAs revealed a functionally coherent and biologically relevant set of regulatory processes. Among the top 20 enriched KEGG pathways, the majority were associated with intracellular signaling cascades (e.g., MAPK, Hippo, and PI3K-Akt pathways)[130] [131] and immune/infection-related mechanisms (e.g., ubiquitin-mediated proteolysis, antigen presentation, and viral infection pathways) [132]. A particularly relevant observation is the enrichment of the *Herpes simplex virus* infection pathway, belonging to the same Herpesviridae family as cytomegalovirus (CMV) [133], the infectious agent in our “Infection” cohort. Notably, as shown in the Upset plot (Figure 6B), this pathway contains the highest number of genes exclusively targeted by the differentially central microRNAs, with no overlap across other enriched pathways. This unique pattern suggests that the Herpes virus response pathway is especially susceptible to regulatory remodeling driven by microRNAs with phenotype-specific centrality changes [134]. Further insights emerge from Table 1, where the mean PageRank values of microRNAs associated with each enriched pathway were compared across phenotypes. A consistent trend is observed: in the rejection phenotype, the average centrality of these microRNAs is reduced compared to both control and infection groups. This decrease, statistically supported by the p-values from pairwise comparisons ($p < 0.001$ in many infections vs. rejection cases), suggests that during acute rejection, the regulatory influence of these microRNAs within key biological pathways is diminished. Their reduced centrality role in the rejection network reflects a disruption or suppression of post-transcriptional control mechanisms. These findings underscore a key insight of our study: although microRNAs are broadly multifunctional and engage in regulatory interactions with multiple targets consistently across all phenotypes, centrality regulation shifts of specific subsets of them exhibit functional consequences. These high-impact microRNAs likely act as “centrality switches” within the multilayer network, playing pivotal roles in guiding the system toward specific phenotypic outcomes. Their altered topological positioning suggests that, beyond serving as concentration-based biomarkers as commonly explored in the literature they may function as topological regulators whose central position within the network enables them to influence the flow of information and coordination among molecular pathways. This perspective aligns with the growing recognition of microRNAs not merely as passive effectors of gene silencing, but as dynamic modulators capable of orchestrating large-scale transcriptomic reprogramming [135] [136]. Beyond the specific findings, an important contribution of this work lies in the methodological framework employed.

From a systems biology perspective, our approach contributes to the broader framework of network medicine, which aims to characterize disease not as a result of individual gene dysfunctions, but as perturbations of entire regulatory modules and their interconnections. The identification of topologically central microRNAs whose role varies across phenotypes supports the idea that pathological states emerge from shifts in regulatory influence rather than from isolated expression changes. The multilayer network model and ensemble centrality estimation provide a robust and scalable strategy for exploring microRNA–mRNA regulatory dynamics [118]. Unlike conventional differential expression analyses that primarily assess expression-level changes, this network-based approach captures the topological and relational complexity of molecular interactions, offering insights into how regulatory elements shape network architecture across conditions [137]. Importantly, the method is data-driven and modular, composed of interchangeable analytical steps, making it adaptable to diverse datasets and applicable regardless of specific biological assumptions on microRNA–mRNA interactions. As such, it offers a generalizable framework for investigating phenotypic interplay and molecular remodeling and may serve as a valuable foundation for future studies aimed at understanding regulatory reprogramming in disease, infection, or therapeutic intervention. We acknowledge that the present study was conducted on a relatively limited cohort, which may restrict the generalizability of the findings. This limitation is partly due to the inherent difficulty in obtaining well-characterized clinical samples from patients undergoing cardiac transplantation with these specific phenotypes [138]. Nevertheless, the consistent topological patterns observed across phenotypes provide a solid foundation for further investigation. Future work will focus on expanding the framework to larger, independent cohorts to enhance statistical robustness and external validity. Additionally, the integration of longitudinal data will be pursued to better capture the temporal dynamics of microRNA–mRNA interactions and model phenotype transitions through evolving network centrality [139]. Importantly, future studies will also aim to include clinical and experimental validation to confirm the regulatory relevance of key microRNAs identified by the network-based approach.

4 Neural Network Modeling of Time Series Data

Part of the PhD has been dedicated to the study and modeling of longitudinal and time series data using deep learning techniques. Within the collaboration with I4 Consulting S.R.L., a recurrent neural network model was developed for the forecasting of streamflow time series in the mountain catchment of Posina. The model's performance was evaluated against traditional approaches, and the results were published in Water (MDPI) in the paper titled "Streamflow Forecasting: A Comparative Analysis of ARIMAX, Rolling Forecasting LSTM Neural Network and Physically Based Models in a Pristine Catchment" [140]. Building on this experience, the same methodological approach was extended to the biomedical domain during a six-month research stay at London South Bank University (LSBU), under the supervision of Professor Enrico Grisan. There, a Variational Autoencoder (VAE) model was developed and validated to analyze time-concentration profiles of tacrolimus, a key immunosuppressant used in organ transplantation. The goal was to identify covariates influencing tacrolimus pharmacokinetics in different patients. This work resulted in the publication "Uncovering Population PK Covariates from VAE-Generated Latent Spaces" [141] presented at IEEE EMBC 2025 and indexed in IEEE Xplore. While belonging to different application domains, both projects rely on deep learning architectures for temporal data analysis. The methodological expertise developed in one setting provided the foundation for successfully applying and adapting these models in heart patients follow-up.

4.1 Streamflow forecasting: A Comparative Analysis of ARIMAX, Rolling Forecasting LSTM Neural Network and Physically Based Models in a Pristine Catchment

4.1.3 Aim

The objective of this study is to evaluate and compare the performance of three distinct streamflow forecasting approaches: ARIMAX, Long Short-Term Memory (LSTM) neural networks with a rolling forecasting training and inference procedure, and a physically based hydrological model (HEC-HMS) applied to a fast-responding Alpine catchment. The analysis is conducted using 13 years of high-resolution (hourly) hydrometeorological data collected from the Posina River basin, a pristine watershed in Northern Italy characterized by strong seasonal variability and extreme rainfall events. The comparison aims to evaluate each model's ability to capture flow patterns and dynamics across a range of hydrological events, with particular emphasis on flood forecasting accuracy and generalization across diverse flow regimes. Performance metrics include Nash–Sutcliffe Efficiency (NSE), Kling–Gupta Efficiency (KGE), and Mean Absolute Error (MAE), evaluated over both the entire test period and selected flood events. Beyond benchmarking, this work also investigates the trade-offs between model complexity, data requirements, and implementation effort. In particular, it explores whether a localized LSTM model trained on single-basin data and deployed in a recursive prediction framework can serve as a scalable and operationally feasible alternative to traditional physically based models, especially in contexts where high-resolution data are available but calibration resources are limited. This case study represents a foundational exploration of neural network-based forecasting methods for temporal data analysis and their practical application compared to physical based mehtods and traditional time series forecasting statistical approaches.

4.1.2 Study Area and Dataset

Study Area

The Posina river is a mountainous river that flows in the North-East of Italy, draining an overall catchment of around 116 km2. Altitude in the catchment ranges between 300 m a.s.l. at the catchment outlet and 2200 m a.s.l. The catchment annually experiences an average precipitation of 1740 mm, with peaks up to 2000–2500 mm. Rainfall distribution is heterogenous across seasons. Highest rainfall volumes concentrate in spring (May) and autumn (October–November), which are the periods when historical flood events take place. Winter rainfalls are usually small. Summer provides relatively small volumes of rainfall but often concentrated in short and intense storm events. The annual average temperature is around 7.5 ◦C, decreasing to minima of −9 ◦C during winter at higher elevations, while increasing to maxima of 16◦C in the summer months. Soil is mainly occupied by forests, with a negligible percentage of catchment occupied by grazing lands, rocks, and small urban areas concentrated in the valleys nearby the network of rivers. Figure 22 provides an overview of the catchment. Soil characteristics were derived from the regional Soil Atlas developed by the Environmental Protection Agency (ARPA Veneto). The Atlas provides estimates for a proper characterization of soils within the Posina catchment, in terms of porosity, surface infiltration capacity, and soil conductivity. A characterization of the rainfall regime and the

corresponding intensity–duration–frequency (IDF) curves for the meteorological stations used in this study is available through ARPAV at https://www.arpa.veneto.it/dati-ambientali/dati-storici/meteo-idro-nivo/precipit-max. The data were retrieved on 15 September 2022.

Dataset

This work takes advantage of a relatively long dataset of hydrologic data provided at an hourly temporal resolution. Meteorological and discharge stations within the catchment and nearby are administered by the Environmental Protection Agency, that makes those data available without any limits of elaboration usage and distribution. The dataset starts in January 2010, and all gauging stations considered are actually in place and still running. Unfortunately, the flow gauge has a long period of no records from 15 July 2022 to 31 July 2023. We thus decided to limit the dataset in correspondence of the summer of 2022, resulting in almost 13 years of hourly precipitation, temperature, and discharge data to work with. Streamflow data are measured very close to the outlet of the Posina river at a place named Stancari, where it flows in the Astico river, and serves as our focal point. Temperature and precipitation are measured in correspondence of six meteorological stations, whose location is described in Figure 22. The distribution of meteorological stations spans across the different areas of the catchment, characterized by different altitudes. An additional meteorological station was considered outside the catchment, as it is the closest to the

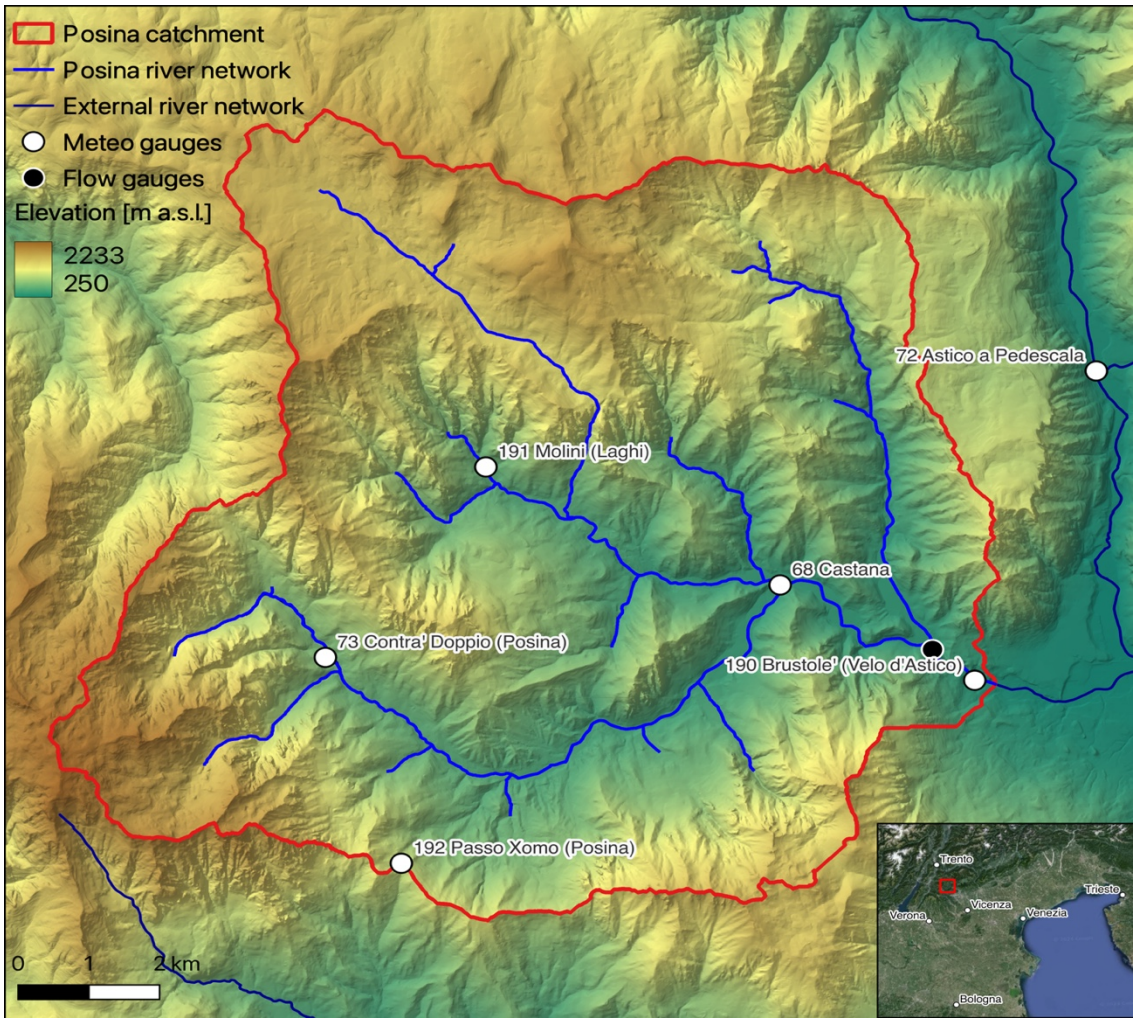
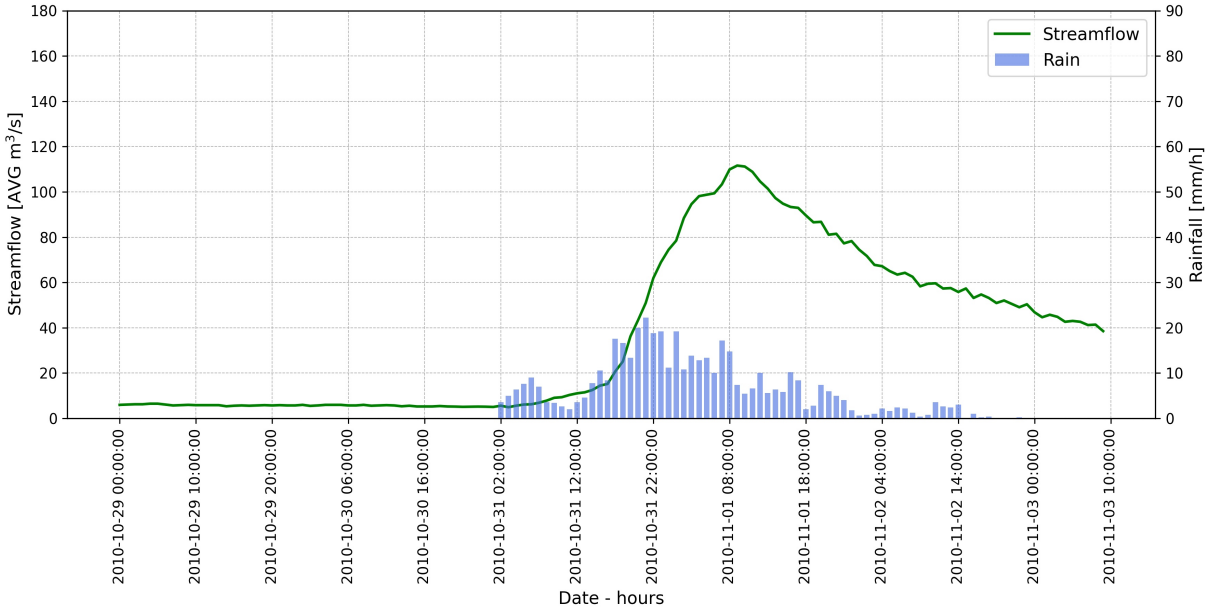
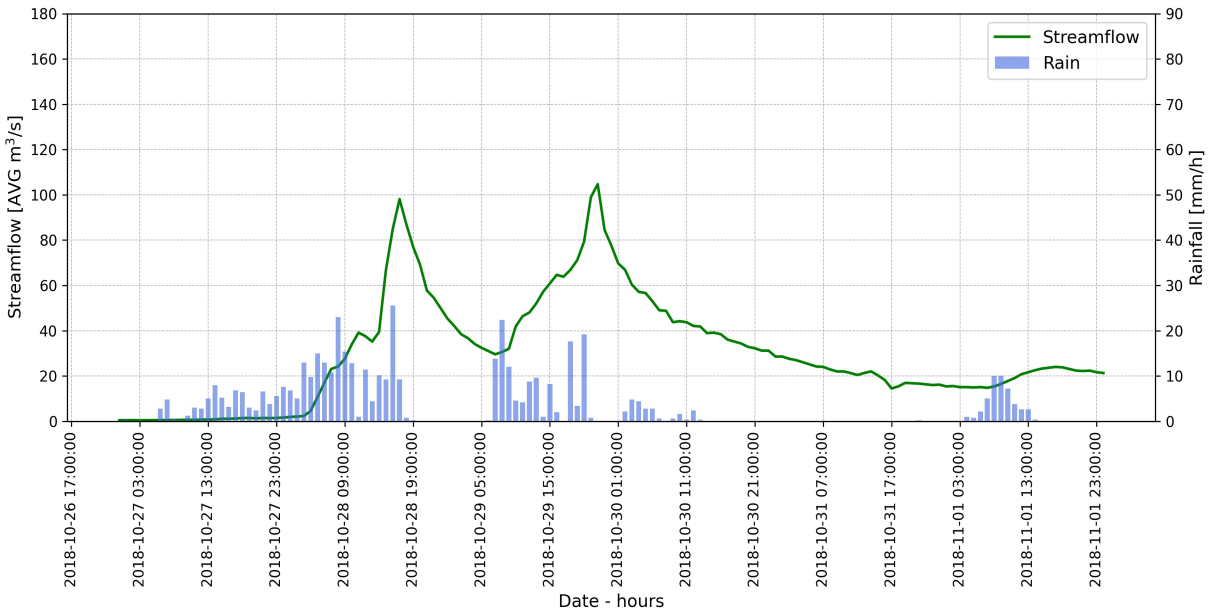


Figure 22. The Posina river network and its catchment.

north-east part of the Posina catchment. Figure 23 a,b show two of the most intense flood events recorded within the monitored period. Both flood events occurred between the end of October and the beginning of November of 2010 and 2018. Figure 24 presents the flow duration of the Posina's catchment, which permits the characterization of the natural hydrologic regime of the basin [142]. The hydrologic regime shows an erratic behavior, typical in a catchment where the mean interarrival between flow-producing rainfall events is larger than the typical duration of resulting flow pulses. A wider range of streamflows is observed between events, and the preferential state of the system is typically lower than the mean [143]. The Posina River may show no discharge in between summer rainfall events.



(a)



(b)

Figure 23. (a) October–November 2010 flood event. (b) October–November 2018 flood event.

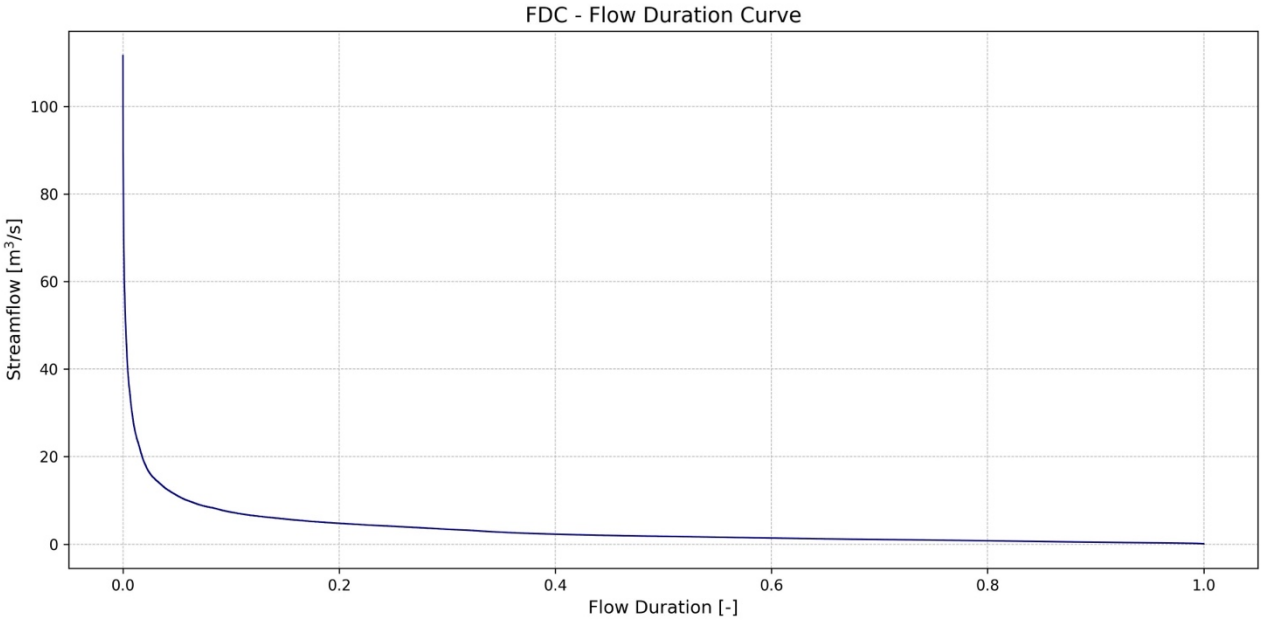


Figure 24. Long-term flow duration curve (FDC) based on historical data from January 2010 to July 2022.

4.1.3 Experimental procedure

Before developing and training the models, we pre-processed the dataset to enhance its quality and eliminate inconsistencies or noise that could affect the accuracy of our models. First, we removed out-of-scale values using prior knowledge of the maximum and minimum hourly streamflow of the catchment area. Next, to address missing values, we applied linear interpolation for nearly all features, except for rain and temperature data, due to their nonlinear behavior. The physically based model, developed using HMS software, automatically handles missing sensor values by discarding them and interpolating them through inverse distance weighting (IDW) using data from other sensors. For the autoregressive and deep learning models, we imputed and pre-processed the temperature and rain-related missing data using the same method. This will allow us to give more importance to the sensors closer to the one presenting the missing value. We have fitted all the models using 70% of the dataset and tested the forecasting capabilities on the remaining 30% except for the HEC-HMS model, where we have included as part of the fitting data also the flood event of October–November 2018 Figure 2b. In order to improve the stability of the LSTM model, we have scaled the data using Min–Max Normalization method. This transformation ensures that the minimum value of each feature becomes 0, the maximum value becomes 1, and all other values are scaled proportionally in between. To emulate the forecasting behavior of the other models with LSTM, we employed a rolling forecasting technique for streamflow prediction. This technique involves a sliding window to iteratively determine streamflow values. Importantly, the model's predictions were used as input to forecast subsequent hours, allowing the model to capture the hydrological processes going on within the watershed, such as streamflow recessions corresponding to dry days. This approach in turn strongly reduces the LSTM model initial conditions importance, as the model quickly becomes independent by the initial observed values provided as it only uses predicted streamflow values after m prediction instances. A visual simplified explanation of this technique can be found in Figure 25, where (h) and m were set as an example to 6 for an easier

visualization. By testing the LSTM model with the rolling window technique, we were able to compare the LSTM model with ARIMAX and the physically based model.

Rolling forecasting example: h = 6, m = 6.	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	00:00	01:00
1° Prediction instance	t−5	t−4	t−3	t−2	t−1	t0	t+1	t+2	t+3	t+4	t+5	t+6				
2° Prediction instance		t−5	t−4	t−3	t−2	t−1	t0	t+1	t+2	t+3	t+4	t+5	t+6			
3° Prediction instance			t−5	t−4	t−3	t−2	t−1	t0	t+1	t+2	t+3	t+4	t+5	t+6		
4° Prediction instance				t−5	t−4	t−3	t−2	t−1	t0	t+1	t+2	t+3	t+4	t+5	t+6	

Figure 25. Rolling forecasting. This image depicts an example of the first four prediction instances of the rolling forecasting procedure that has been implemented. h refers to the previous hours of streamflow values used as input (cells with a red border), while m is the number of instances predicted at each inference (cells in dark and light blue). The green cells correspond to real measured values. The orange ones refer to predicted values used as input. The inference procedure moves as a sliding window. Therefore, at some point, the model will use only predicted values as input to compute the forecast for the next m values. The final forecast is obtained by combining the dark blue predicted values.

ARIMAX model

The ARIMA model is a well-known autoregressive model used for time series forecasting in various contexts. In recent years, it has garnered significant attention in hydrological studies. For this study, we used the extension of this model called ARIMAX, which enables the inclusion of exogenous variables. Models like ARIMAX rely on the assumption of stationarity of the time series under analysis. This does not imply that the series remains unchanged over time, but rather that the way it changes remains consistent over time. To ensure the suitability of ARIMAX modeling, we assessed the stationarity of the streamflow target time series for the Posina catchment using the augmented Dickey–Fuller (ADF) test [144]. The calculated p-value was 8.90 ×10−29, confirming that the time series is stationary. The ARIMAX model relies on the following parameters: *p*, representing the autoregressive (AR) component or the number of lag observations in the model; *d*, denoting the degree of differencing required to achieve stationarity in the time series; and q corresponding to the size of the moving average (MA) window. Given that the streamflow time series passed the augmented Dickey–Fuller (ADF) test for stationarity, the parameter d was set to 0. To determine the values of parameters p and q, we analyzed the autocorrelation (ACF) and partial autocorrelation (PACF) plots to determine the optimal *p*, *d*, and *q* values. A parameter optimization analysis was conducted by comparing the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) across different combinations of autoregressive (*p*) and moving average (*q*) orders. These metrics were used to evaluate the trade-off between model fit and complexity, supporting the selection of the most parsimonious ARIMAX configuration. These values have also been refined by exploiting the automatic algorithm presented by Rob J. Hyndman and Yeasmin Khandakar in the work “Automatic Time Series Forecasting: The forecast Package for R” [145], which is a step-wise procedure traversing the space of models efficiently. The ARIMAX model parameters *p*, *d*, *q* have been set to 1, 0, 2, respectively. We have chosen to use the extended version of ARIMA, which allows for the inclusion of exogenous variables (the “X” in ARIMAX),

we were able to incorporate rain and temperature forecasts as additional inputs for the streamflow predictions. Since ARIMAX is an autoregressive model, the forecasting of theionstreamflow involves employing a linear combination of past values of the streamflow along with any exogenous variables. The general ARIMAX equation is depicted in (4 where ϕ is the autoregressive term and p is the relative order. y_t and y_{t-i}, are, respectively, the dependent variable value at time t and the lagged values of the dependent variable. θ is the moving average term with the relative order q. r represents the set of exogenous variables. $X_k t$ is the exogenous feature k at time t. βk and ϵ are, respectively, the strength coefficient of the exogenous variables and a random noise between the observed and predicted value.

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{r} \beta_k X_{kt} + \epsilon_t \quad (4)$$

Long Short-Term Memory (LSTM) Recurrent Neural Network

The Long Short-Term Memory (LSTM) model is a specific type of recurrent neural network (RNN) that has received a lot of interest in the hydrological field [146,147]. It has been presented by Hochreiter and [148], to overcome the limitation of capturing and remembering long-term dependencies in sequential data. Before the training procedure, the constructed dataset was pre-processed to improve the efficiency and generalization of the model [149] as explained in the experimental setup section. Hyperparameter tuning plays a critical role in shaping the learning dynamics of the LSTM model and directly influences its forecasting performance. To identify an optimal configuration, we conducted a manual tuning procedure, systematically testing a range of hyperparameter combinations. The tuning process focused on the following key parameters: the model complexity (i.e., number of hidden layers and number of LSTM units per layer), batch size, initial learning rate, and the temporal structure of the input–output mapping, defined by the number of input hours (h) and predicted hours (m) at each inference step. The model consists of eight LSTM layers comprising 100 units each. A simplified representation of the model architecture is shown in Figure 26. The learning rate was initially set to 1 ×10−3 and dynamically adjusted during training to ensure adequate convergence and generalization. As a loss function, we employed the Mean Absolute Error (MAE) [150], described in the equation (7 where n denotes the total number of instances, yi denotes the observed streamflow value, and $\hat{y}_i$ denotes the corresponding prediction. To prevent overfitting and improve generalization, an early stopping mechanism was implemented: if the validation loss did not improve for 25 consecutive epochs, the training process was halted and the model weights corresponding to the lowest validation loss were retained. Particular attention was given to exploring the effect of batch size and input sequence length (h). The sensitivity analysis was conducted to assess the impact of varying the batch size and number of input hours (h) on model training with a dynamically adjusted learning rate, and the results are reported in the Supplementary Materials. Among the tested configurations, the final model was selected based on the best trade-off between training loss and validation loss, ensuring both the fit quality and generalization capability. The model was implemented using Python programming language, alongside the machine learning libraries Keras and TensorFlow. Each input to the trained model

comprised all recorded streamflow, rainfall, and temperature values for the preceding *h* hours leading up to the inference instance. Additionally, the model incorporated the recorded hourly rainfall and temperatures for the next *m* hours forecast as for the exogenous values of the ARIMAX model. We assume that only the streamflow is recurrently predicted, whereas the weather data are the real ones (not predicted) at each time point. After tuning the parameters and hyperparameter, the optimal values for *h* (previous hours) and *m* (number of forecast streamflow value at each instance) were determined to be 24, resulting in the model using the last 24h of data to predict the subsequent 24h.

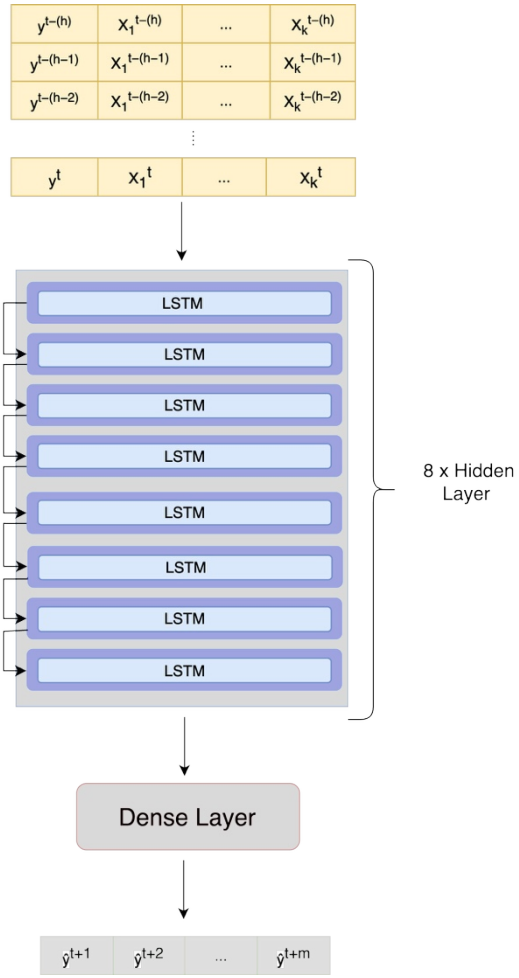


Figure 26. LSTM model configuration. Yellow boxes show the input $y^{(t-h)}$ streamflow values observed before the inference instance. X_1 and X_k are the additional features used for the preceding h hours (streamflow, rainfall, and temperature) and the prediction of rainfall and temperature for the next m hours. $\hat{y}$ reflects the streamflow values predicted in the subsequent m hours.

Physical Based Model

Models presented in previous paragraphs were compared with a physically based model which was developed with the software Hydrologic Modelling System (HMS), version 6.11, by the US Army Corps of Engineers Hydrologic Engineering Center (HEC). HEC-HMS allows users to combine

various procedures implemented to reproduce hydrological processes and build a customized hydrological model. In the present application, the hydrological model was set to allow continuous simulations. This choice increases the complexity of the model as several additional processes must be considered than in event-based models, such as evapo-transpiration and soil moisture accounting, for example. The spatial distribution of many variables influencing the runoff separation process implies the adoption of a spatial scale smaller than the entire watershed to account for such variability (e.g., elevation within the basin, land use, or soil type). The Posina river basin was thus discretized in 34 subcatchments (Figure S1 in Supplementary Materials) with an average size of 3.45 km2, within which spatial variables affecting hydrologic processes can be considered homogeneous. The number of subcatchments was not exaggerated to avoid excessive model complexity. In the proposed hydrological modeling, the soil moisture accounting model (hereinafter SMA) was adopted as it allows for describing the long-term dynamics of water content [150,151]. The SMA model represents the watershed with a series of interconnected storage layers. Current storage contents are calculated during the simulation and vary continuously both during dry or wet periods. Full details of this methodology can be found in the HEC-HMS Technical Reference Manual [152], from which a brief explanation was taken and described below. The different storage layers considered in this application are given as follows:

- Canopy-interception storage represents the precipitation that is captured on trees, shrubs, and grasses, and does not reach the soil surface. Precipitation is the only inflow. Water in canopy interception storage is removed by evaporation.

- Surface-interception storage is the volume of water held in shallow surface depressions. Inflows come from precipitation not captured by canopy interception and in excess of the infiltration rate. Outflows can be due to infiltration and to evapotranspiration (ET).

- Soil-profile storage represents the water stored in the top layer of the soil. Inflow is infiltration from the surface. Outflows include percolation to a groundwater layer and ET. The soil profile is subdivided into two distinct layers. The upper zone is defined as the portion of the soil profile that will lose water to ET and/or percolation. The tension zone is defined as the area that will lose water to ET only. The upper zone represents the water held in the pores of the soil. The tension zone represents the water attached to soil particles. ET occurs from the upper zone first and tension zone last.

- Groundwater storage layers in the SMA represent horizontal interflow processes. In this application, only one groundwater layer was used. Water percolates into groundwater storage layers from the soil profile. Losses from the groundwater storage layer are due to groundwater flow or to deep percolation. In the latter case, this water is consider lost from the system.

Details on the computation of flows between interconnected storage layers can be found in the HEC-HMS Technical Reference Manual [152]. Evapotranspiration was calculated at an hourly timescale through the Hargreaves method which also allow one to estimate the shortwave radiation from temperature data [153,154]. The hydrological model is completed with modules describing

how the model transforms excess precipitation into runoff (transform process) and how the water content stored within the soil and groundwater layer becomes baseflow runoff (baseflow process). The Transform process to route excess precipitation to the subcatchment outlet was described with the Clark Unit Hydrograph Model [155]. This method explicitly represents two critical processes in the transformation of excess precipitation to runoff: (1) the translation (or movement) of excess precipitation from its origin throughout the watershed to the outlet; and (2) the attenuation (or reduction) of the magnitude of the discharge as the excess precipitation is temporarily stored throughout the watershed. The baseflow process is described by the Linear Reservoir Model, which uses one linear reservoir to simulate the recession of the baseflow after a storm event. According to the model, baseflow magnitude linearly depends on the amount of water stored within the groundwater layer of the SMA model. The linear release of water can be repeated in a waterfall process to increase the baseflow attenuation. The methods described above provide a full description of the processes occurring at subcatchments which determine how rainfall forcing the system becomes discharges input within the river network. Flows are calculated at the subcatchment outlet and require an additional method to describe what occurs within the channels up to the catchment outlet. The Posina river has a typical mountainous behavior with steep and narrow channels, and as such, any attenuation process can be considered negligible. The methodology adopted for describing the routing process within the river network is the Lag Model, which simply considers that there is a translation of discharge within the river without any attenuation. Translation depends on the travel time within each river reach, which in turn mainly depends on the reach length and slope. The hydrological model composed as described above requires the definition of several parameters for each subcatchment and river reach characterizing the Posina catchment. A list of the required parameters is provided below.

- Canopy Interception
    - Canopy storage [mm], represents the maximum amount of water that can be held on leaves before through-fall to the surface begins.

- Surface Interception
    - Surface storage [mm] represents the maximum amount of water that can be heldon the soil surface before surface runoff begins.

- Soil Moisture Accounting
    - Soil storage [mm], total storage available in the soil layer;
    - Tension storage [mm], amount of soil storage that is not drained by percolation but only by evapotranspiration;
    - Groundwater storage [mm] represents the total storage in the groundwater layer;
    - Impervious percentage [%], percentage of the subcatchment with direct runoff production (no infiltration);
    - Maximum infiltration rate [mm/h] sets the upper bound on infiltration from the surface storage into the soil;
    - Soil percolation [mm/h] sets the upper bound on percolation from the soil storage into the groundwater;

    - Groundwater percolation rate [mm/h] sets the upper bound on deep percolation.

- Clark Unit Hydrograph
    - Time of concentration [h] defines the maximum response time in the sub-basin;
    - Storage coefficient [h], accounts for storage effects within the subcatchment surface.

- Linear Reservoir Baseflow
    - Groundwater coefficient [h] is used as the time lag on a linear reservoir for transforming water in storage to become lateral outflow;
    - Number of steps [−], which increases the attenuation of baseflow (minimum attenuation with a single step; attenuation increases as the reservoir release is repeated several times).

- Reach Routing
    - Lag [min], time that the inflow hydrograph will be translated.

All parameters listed above were estimated depending on information available at the Posina catchment (soil atlas, soil use) and on the physical features of each element of the river system (reach length, slope and width). Some of these were determined thanks to empirical formulation (time of concentration), while others were obtained by the calibration of the model against observed flows. The model parameters have been calibrated on the same training period of the other models. A single set of parameters was thus adapted. The calibration process focused on a limited set of parameters. Specifically, canopy and surface storage values were calibrated uniformly across the basin within plausible ranges (0–30 mm and 0–50 mm, respectively), reflecting the catchment's forested land cover. The runoff response was shaped by calibrating the storage and groundwater coefficients of the Clark unit hydrograph and baseflow modules, expressed as multipliers of the time of concentration (Tc). The final values, 16×Tc and 170×Tc, respectively, were selected after iterative tuning to match the observed flood dynamics. The model also requires initial conditions which, however, have had negligible impact on the model performance, as in long-term continuous simulations, the initial error is forgotten by the system after a rainfall determining soil saturation.

4.1.4 Validation Metrics

All predictions generated by the models have been evaluated and compared using multiple performance criteria. We utilized two well-known metrics widely applied in the hydrological field: the Nash–Sutcliffe Efficiency index (NSE) and the Kling–Gupta Efficiency index (KGE). Additionally, we have measured the previously presented Mean Absolute Error (MAE) from Eq. 4, not only as a loss function for optimization but also as an evaluation criterion. To measure MAE we have exploited the method implemented into the scikit-learn python library [90]. To compute NSE and KGE we have used the hydroeval open source python library [156].

Nash–Sutcliffe efficiency index (NSE)

The Nash–Sutcliffe efficiency index is a widely used and reliable statistic for assessing the goodness-of-fit of hydrologic models. One of its main advantage is that it can be applied to a variety of model types [157].
In (5, y_t is the observed value at time t, $\hat{y}_t$ is the predicted value at the same time t, $\bar{y}_t$ is the mean of the observed values, and n refers to the total number of observations.
The NSE value ranges from $-\infty$ to 1, where 1 signifies a perfect model with an estimation error variance from the original measurements equal to zero. An NSE value equal to 0 indicates that the model under investigation incorporates the same predictive capabilities as the mean of the forecast time series. Higher NSE values are associated with greater predictive capabilities.

$$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} \tag{5}$$

Kling–Gupta Efficiency Index (KGE)

The Kling–Gupta efficiency index (KGE) [158] was introduced as an improvement over the NSE, in order to capture an additional aspect of the model such as correlation, bias, and variability. The term r in (6 represent the Pearson correlation coefficient between observed and simulated values. α is the measure of the flow variability error calculated as the ratio between the variance of the simulated time series and the variance of the observed time series. β is the ratio of the mean of the simulated values to the mean of the observed values, also identified as the bias term. Analogous to NSE, KGE = 1 indicates perfect agreement between simulations and observations while KGE = 0 or KGE < 0 indicates that the mean of observations provides better estimates than simulations [159,160].

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \tag{6}$$

Mean Absolute Error (MAE)

The mean absolute error (MAE) described by the (7. The main difference of measuring the predictive capabilities of a model with MAE instead of using NSE or KGE, providing a straightforward measure of the absolute accuracy of the model's predictions [161]. The optimal MAE value is 0, indicating perfect prediction accuracy. Higher values of this index are associated with worse forecasting capabilities. Since we are computing the absolute difference between the observed and the predicted value, the MAE unit measure is the average (AVG) m^3/s measure at each hour.

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \tag{7}$$

4.1.5 Test Set Forecasting Results

We evaluated the forecasting performance of all models over the entire testing set by computing the metrics previously described, namely NSE, KGE, and MAE. These metrics were used to assess each model's predictive accuracy on unseen testing data, and the corresponding results are summarized in Table 8. To provide a visual comparison, we present plots showing predicted versus observed streamflow values. Each plot reports time on the x axis (in dates) and streamflow on the y axis, expressed in average m3/s. In addition to the standard linear scale, we also include plots using the symlog (symmetric logarithmic) scale on the y axis. This scale is particularly useful for representing data that spans several orders of magnitude while retaining sensitivity to low-flow conditions, including near-zero and negative values, without distortion.

Metric	ARIMAX	LSTM	Physical Based Model
NSE [-]	0.67	**0.93**	0.82
KGE [-]	0.50	0.82	**0.85**
MAE [AVG m^3/s]	1.16	**0.75**	1.27

Table 8 Models forecasting metrics performances on the testing set.

ARIMAX Forecasting Results

The images in Figure 27 a,b depict the forecasting results of the ARIMAX model on the testing dataset, using the autoregressive principle discussed in the preceding section. The green line refers to the real observed values, while the red line refers to the predicted simulated values from the model. Measured values of NSE, MSE, and MAE has been collected in Table 8. The NSE stands at 0.67, indicating that the model's estimation error variance is lower compared to the variance derived from the mean observed data. However, the KGE value hovers around 0.50 highlighting the imperfect correlation between the observed and predicted values. The average magnitude of errors between predicted and actual values, as measured by the mean absolute error, is 1.16 m^3/s.

LSTM Forecasting Results

As mentioned earlier, we assessed the predictions produced by our rolling forecasting LSTM model. To do this, we reconstructed the entire testing set by forecasting the next 24 h, advancing one hour at a time in the inference process. We retained only the value forecast for the next hour (t+1) to reconstruct the entire testing set. At each step, we discarded the forecast values from t+2 to t+48. In this way, the model quickly became independent from the initial condition, relying on its own predictions to forecast the subsequent values. Figure 28 a,b depicts the LSTM model forecasting. For simplicity, the color of the observed and predicted values is again green and red, respectively. The evaluation metrics values presented in Table 8 reveal that the LSTM model demonstrates proficient predictive capability for the overall testing set, with NSE, KGE, and MAE values of 0.93, 0.82, and 0.75, respectively.

Physical Based Model Forecasting Results

For the same period of testing, we evaluated the prediction capability of the traditional physically based hydrological model. Figure 29 a,b shows the physical based model forecast on the test period. As for the previous images, the green line refers to the original observed values, while the red line refers to the predicted values from the model. From Table 8, we can see that the physically based model achieved NSE, KGE, and MAE values of 0.82, 0.85, and 1.27, respectively.

4.1.6 Models Performance Comparison During Significant Flood Events

To gain deeper insights into the performance of different models, we investigate their streamflow forecasting capabilities during significant flood events. Specifically, we examine three distinct streamflow conditions available in our dataset, focusing on two significant flood events from 2010 and 2018 and one occasional summer rain event that occurred in July 2021. We selected these three events because they represent a diverse range of hydrological conditions, major floods and a more typical rainfall event, allowing us to test the models' robustness across different scenarios. This diverse selection helps to ensure a comprehensive evaluation of model capabilities under varying streamflow patterns. In all the images, we represent the streamflow observed values in green, while the red, purple, and yellow refer to the LSTM, ARIMAX, and physically based models forecast values, respectively.

Model Comparison Flood Event October–November 2018

The image in Figure 30 (a) depicts the prediction capabilities of the flood event occurs within the validation set of our dataset, specifically between October and November of 2018, of the three models. The ARIMAX and LSTM models were not trained on this particular event, whereas the traditional physically based model included this event as part of its fitting data. During the flood event, two main streamflow peaks were observed. The first peak occurred on October 28th around 16:00, reaching approximately 100 AVG m^3/s, while the second peak was recorded on October 29th at approximately 22:00, with a maximum value close to 105 AVG m3/s. Both ARIMAX and LSTM models showed an underestimation of the observed values. The LSTM model predicted the two peaks with values of approximately 55 AVG m^3/s and 62 AVG m^3/s, respectively. ARIMAX estimated both peaks at approximately 30 AVG m^3/s. The HEC-HMS physically based model estimated the peaks at approximately 85 AVG m^3/s and 78 AVG m^3/s, respectively, which were closer in magnitude to the observed values. All models showed a generally coherent timing of the peak occurrences, with minimal time shifts relative to the observed values.

Model Comparison Flood Event October–November 2010

The Figure 30 (b) illustrates the 3 different model estimations of the October–November 2010 flood event, which is part of the training set. Including this event in the comparison serves to highlight the behavior of the models on data also used during training or calibration. This provides useful insights into their predictive tendencies, such as overfitting or smoothing. During the 2010 flood event, a single pronounced streamflow peak was observed on November 1st around 02:00, reaching approximately 110 AVG m^3/s. The LSTM and ARIMAX models predict peak values very close to the observed data, with estimated maxima of approximately 108 AVG m^3/s and 110 AVG m^3/s, respectively. However, while ARIMAX reproduces the signal with high accuracy due to having been fit on these data, the LSTM model yields a smoother predictive curve. In contrast, the HEC-HMS physically based model significantly overestimated the peak, with a value around 150 AVG m^3/s, and exhibited a slower recession phase compared to both the observed data and the other models.

Model Comparison Occasional Rain Event July 2021

The Figure 30 (c) shows the performance of the three models during the isolated summer rainfall event of 2021. In this case, the streamflow response exhibited a very sharp rise compared to the other events, with the observed peak reaching approximately 35 AVG m^3/s. The HEC-HMS model overestimated the peak with a predicted value of about 42 AVG m^3/s, and showed a slower recession phase relative to the observed streamflow. Both LSTM and ARIMAX models underestimated the peak, with predicted values below 20 AVG m^3/s. The LSTM model, in particular, produced a smoother streamflow response compared to the observed rapid dynamics.

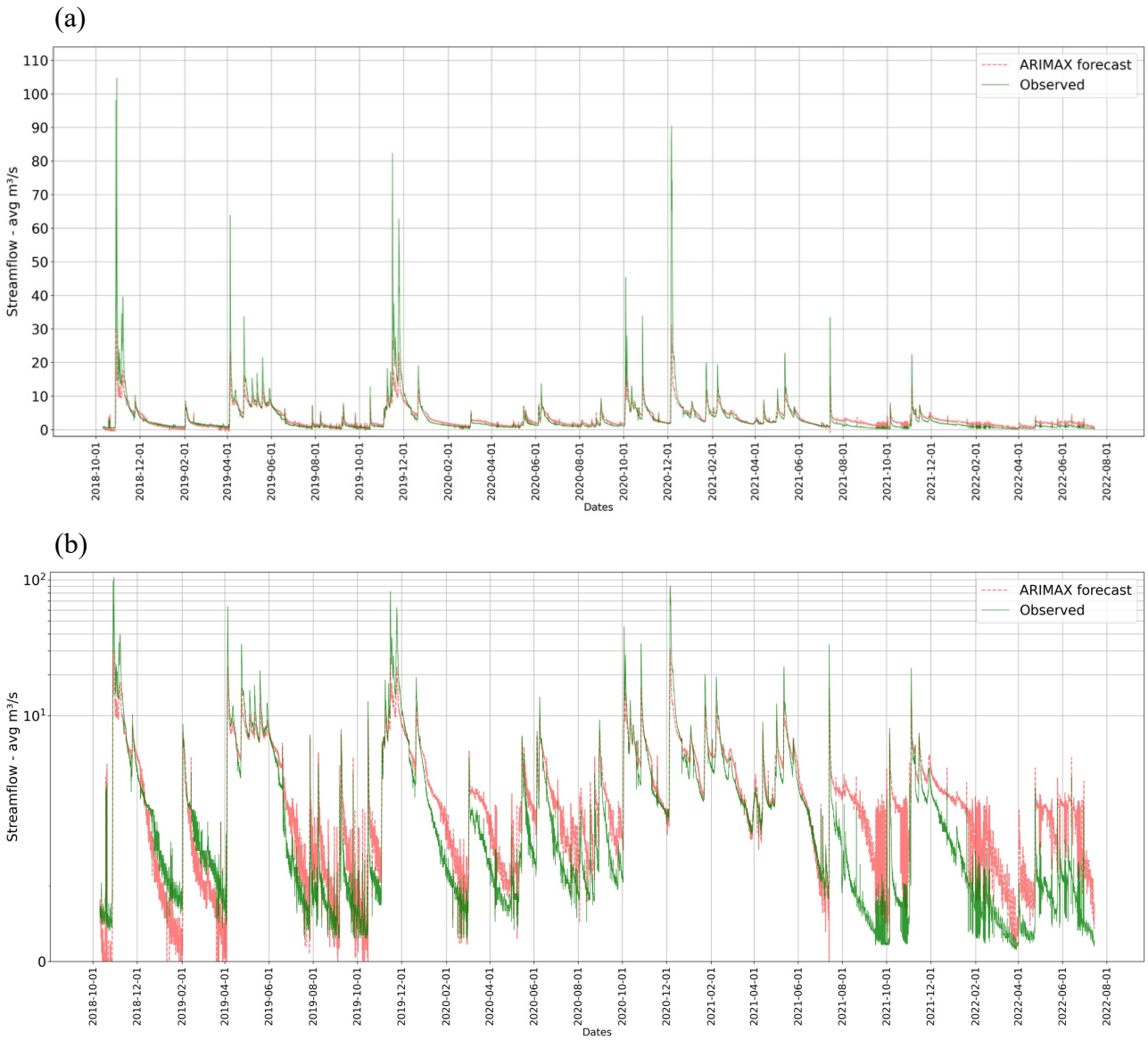


Figure 27. (a) ARIMAX testing set hourly forecasting. In red, the predicted values from the ARIMAX model, and in green, the original observed values for the same period. (b) ARIMAX testing set hourly forecasting. y axis symlog scaled. The y axis is scaled using the symlog (symmetric log) scale method.

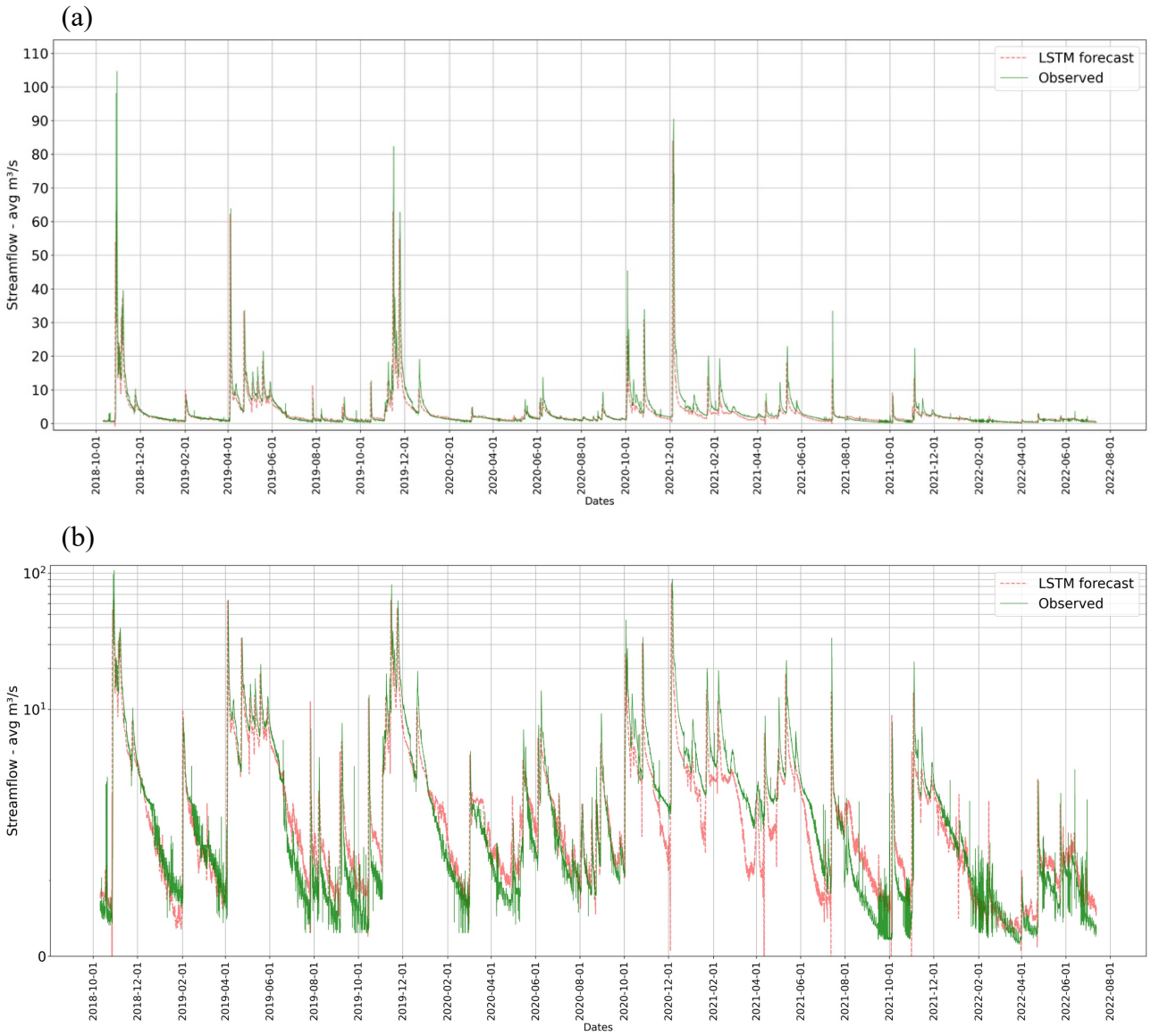


Figure 28. (a) LSTM testing set rolling forecasting. The red line depicts the predicted values of the LSTM model, while the green line represents the original observed values. (b) LSTM testing set rolling forecasting, y axis symlog scaled. The y axis is scaled using the symlog (symmetric log) scale method.

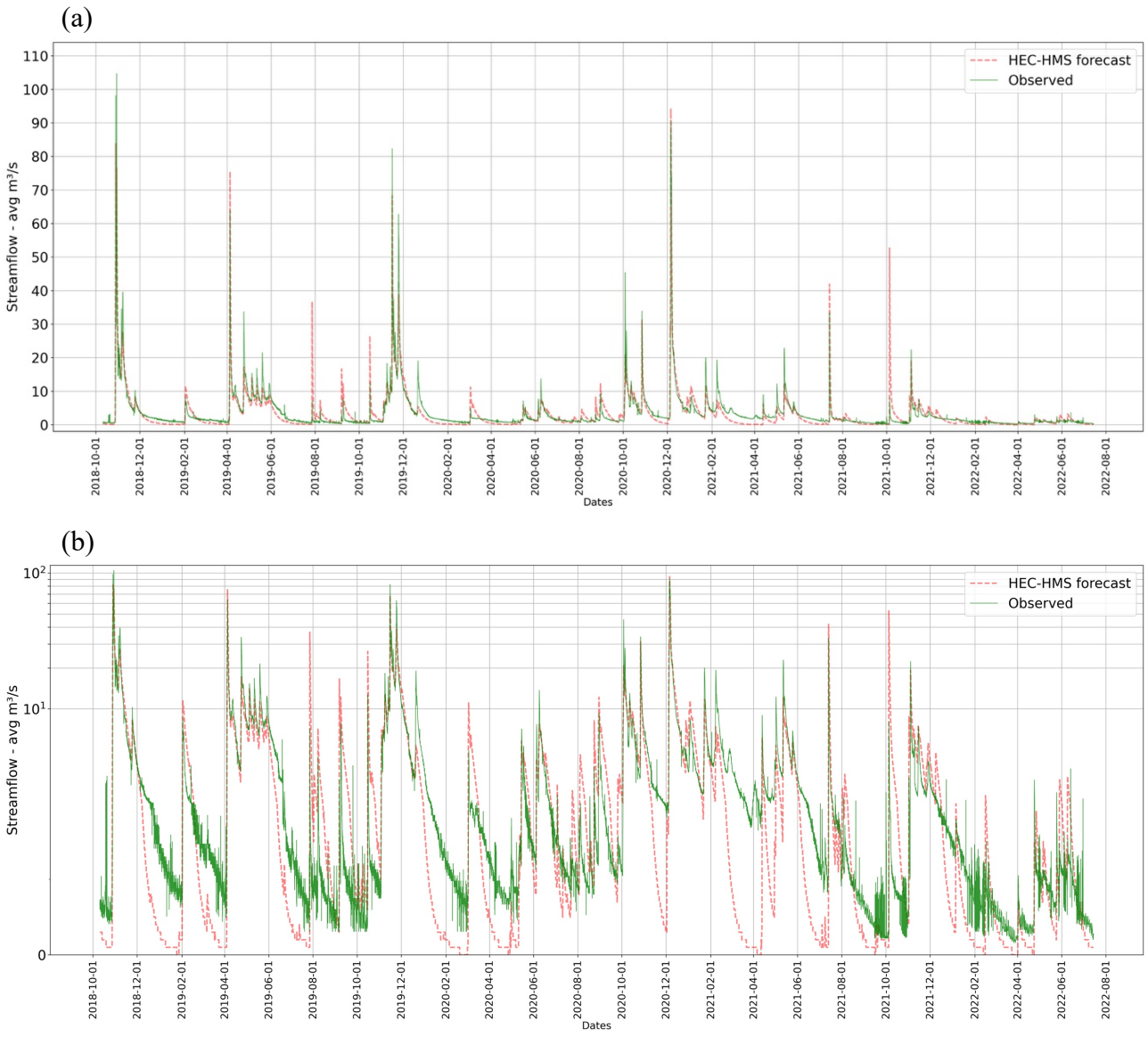


Figure 29. (a) HEC-HMS physically based model prediction on testing set. The red line depicts the predicted values from the physically based model, while the green line shows the original observed values. (b) HEC-HMS physically based model prediction on testing set. y axis symlog scaled. The red line depicts the predicted values from the physically based model, and the green line represents the original observed values.

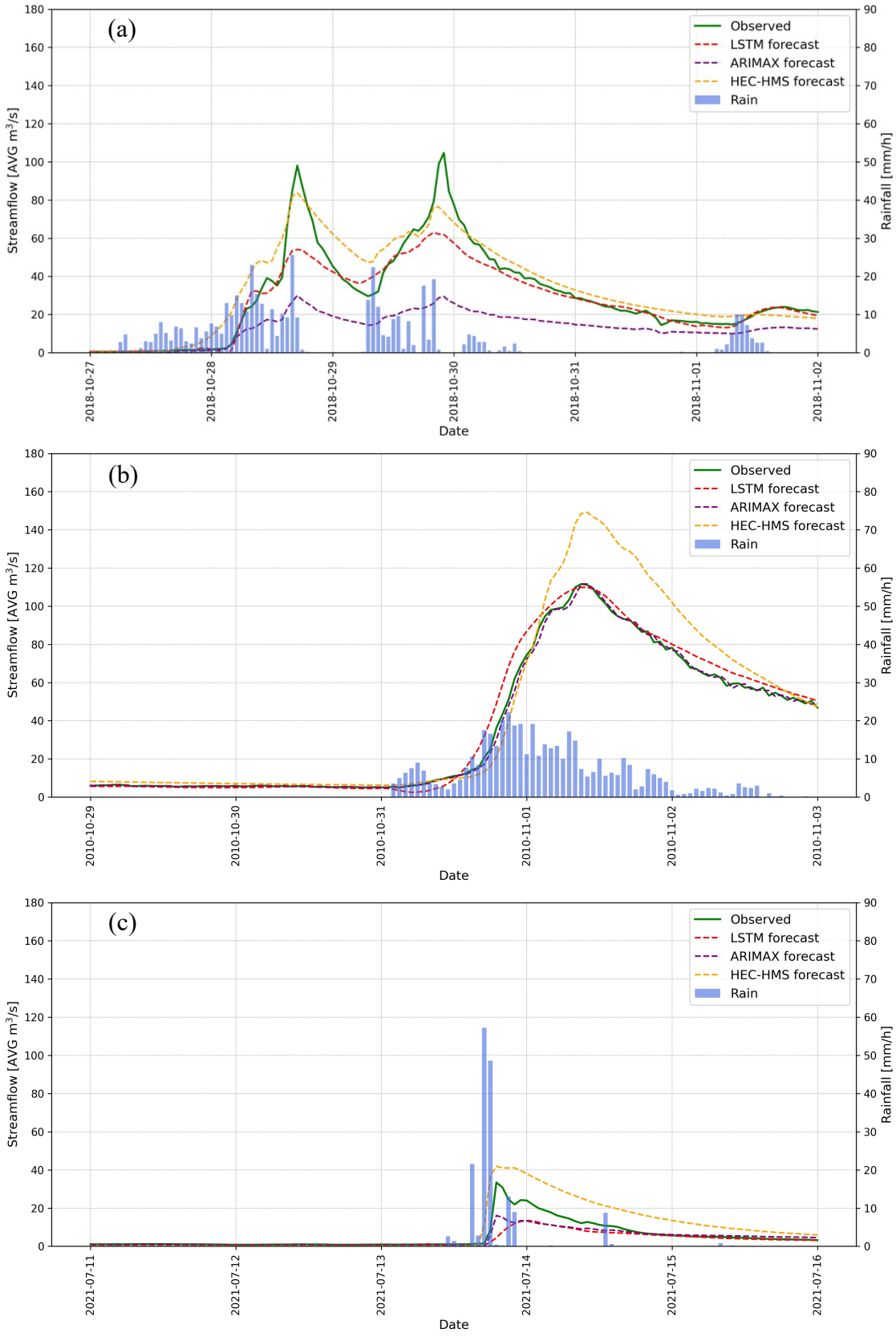


Figure 30. Model performance comparison for (a) October–November 2018 flood event. (b) October–November 2010 flood event. (c) Occasional summer rain event of July 2021.

4.1.7 Discussion: Comparative Analysis of ARIMAX, Rolling Forecasting LSTM, and Physically Based Models in Streamflow Prediction

The In this study, we have compared three different approaches for streamflow forecasting in a challenging Italian natural catchment, characterized by multiple types of soils, intense rainfall periods, and strong seasonal variability. The LSTM and traditional physically based hydrological models have remarkably better capabilities on streamflow forecasting compared to the ARIMAX autoregressive model [140]. Inspecting the images depicted in Figure 27 and reported metrics in Table 8. It is possible to notice that the autoregressive ARIMAX model is capable of catching the overall streamflow trend with sufficient correctness to detect the different streamflow scenarios. However, it struggles with accurately predicting abrupt changes in streamflow, underestimating peaks, and introducing noise over time due to autoregressive prediction behavior [162][163]. Furthermore, the model introduces noticeable noise in regions near zero streamflow values, and as time progresses, the model seems to accumulate errors, leading to a degradation in prediction accuracy over time. Its performance is significantly lower compared to the LSTM and physically based models in capturing peak values and the dynamics of rapid increases. Being simple to implement due to the low computational cost and the well-known multiple programming libraries available for development, it can be used as a baseline approach to study the attitude of the streamflow under investigation. The LSTM model exhibits excellent performance, as evidenced by the metrics reported in Table 1. NSE and KGE values of 0.93 and 0.82, respectively, highlight the model's strength in capturing the correlation between observed and predicted values, indicating good agreement in the temporal pattern of the data [164]. Compared to the estimation of the ARIMAX model (Figure 27) LSTM introduces significantly less noise near zero-flow values, as illustrated in Figure 28 a,b. Furthermore, the LSTM model operates in a rolling forecast mode, where predicted values are fed back as input for subsequent time steps. This recursive prediction strategy demonstrates the model's ability to generalize and maintain stability over time [165], effectively capturing the overall trend of the testing set without external correction. Specifically, it effectively manages most streamflow peaks, even in instances where extreme cases were absent from the training data. The LSTM model shows optimal capabilities in capturing both peak flows and their subsequent recession phases compared to the physically based and ARIMAX models (Figure 29 b, Figure 27 b). This highlights its enhanced ability to handle rapid hydrological responses typical of fast-responding catchments [166], delivering sharper and more stable predictions during high-flow conditions. While the training process of the LSTM model can be relatively time-consuming and requires careful parameter and hyperparameter tuning, its overall computational demand remains considerably lower than that of setting up and calibrating a full physically based model [164]. The LSTM model tends to smooth out slightly rapid streamflow changes, but the model's performance is commendable for streamflow prediction, especially when sufficient training data are available. The physically based model relies on the physical natural behaviors of the catchment which permits an efficient ability to determine the overall streamflow trend, pikes, and rare events that even the other data-driven model may struggle to detect [167,168]. Despite this, the model shows a tendency to overestimate peak values and often exhibited a rapid decline after peak events, missing the smooth recession observed in actual data (Figure 29 b). While the physically based model benefits from its physical basis and ability to incorporate various hydrological processes, it requires careful calibration and may not always generalize well to all

events observed within the training and testing sets. These behaviors are also reflected in the analysis of the three representative flood events illustrated in Figure 30. In the 2010 flood event (Figure 30 b), all models captured the peak timing consistently, with the LSTM and ARIMAX models predicting peak values close to the observed 110 AVG m^3/s. However, the HEC-HMS model significantly overestimated the peak and showed a slower post-peak recession. A similar overestimation by the physically based model is observed during the occasional summer rainfall event of July 2021 (Figure 30 c). The particularly high accuracy of the ARIMAX model in this case can be attributed to the fact that this flood event is included in the training set, allowing the model to directly fit the observed dynamics. In contrast, although the LSTM model was also trained on the same period and demonstrates good performance, it employs a rolling forecasting strategy, where its own previous outputs are iteratively used as inputs for subsequent predictions. This approach introduces a degree of compounding uncertainty, which tends to smooth the resulting predictions. In the 2018 flood event (Figure 30 a), HEC-HMS produced peak estimates closer to the observed values, although with smoother transitions. This behavior is likely due to the incorporation of physical knowledge calibrated for the basin, combined with the fact that the event data were part of the model's calibration set [168]. As discussed in the Results section, both the LSTM and ARIMAX models underestimate peak values in certain cases, likely due to their reliance on previously observed data, which results in smoother predictions and reduced sensitivity to sudden flow changes. However, the LSTM model consistently provides more accurate and homogeneous streamflow predictions compared to the ARIMAX model. It has also demonstrated a remarkable ability to handle different hydrological conditions, often achieving performances comparable to, or even exceeding, those of the physically based model. Being trained specifically on the basin under examination, the LSTM model effectively leverages the rolling forecasting approach, using its own predictions iteratively as inputs for future time steps. This characteristic enables the model to maintain high predictive performance over extended forecast horizons. Furthermore, its adaptive nature and reduced need for complex parametrization make it a practical and efficient alternative for real-world streamflow forecasting applications, especially when compared to the more resource-intensive setup and calibration required by traditional physically based models. Despite the promising results achieved, some limitations must be acknowledged. The LSTM model, implemented with a rolling forecasting strategy, is well suited to adapt when trained on a single basin; however, this study focuses on a single catchment with specific hydrological and geomorphological characteristics, which may limit the generalizability of the findings. The Posina basin is a fast-responding alpine catchment characterized by intense rainfall and strong seasonality, and the trained models may have implicitly learned behaviors unique to this context. As a result, further validation across multiple catchments with diverse hydrological regimes is essential to assess the robustness and transferability of the proposed approaches. Additionally, while the LSTM model demonstrated excellent predictive performance, it remains highly dependent on the availability and quality of historical data [169]. Limited data length, or shifts in climate or land use, could impact the model's effectiveness in real-world applications. Future work will explore the impact of limited data availability and the development of hybrid modeling approaches that integrate the strengths of data-driven and physically based paradigms. Embedding physical constraints or process-based layers within deep learning architectures [170] may improve model interpretability, reduce calibration requirements, and enhance generalization to ungauged or poorly instrumented basins.

4.2 Uncovering Population PK Covariates from VAE-Generated Latent Spaces.

4.2.3 Modeling Purpose: Tacrolimus and Covariate Selection

Building on the expertise developed in time series analysis and the implementation of deep learning models for forecasting complex, dynamic systems, we adapted these methods to pharmacokinetic data of immunosuppressant drugs used in transplant recipients. As in streamflow forecasting, where neural networks effectively captured non-linear temporal dependencies, here they were applied to tacrolimus concentration–time profiles to identify latent patterns and clinically relevant covariates influencing immunosuppressant absorption and disposition. Pharmacokinetic (PK) aims to understand the absorption, distribution, metabolism, and excretion of drugs within the body. Population pharmacokinetic (PopPK) modelling plays a crucial role in understanding drug behaviour across diverse patient populations, allowing for the optimization of individualized dosing regimens. Covariates such as age, weight, ethnicity, genetic factors and others often influence drug behaviour. With the growing emphasis on personalized treatment, reliable models are essential for identifying key covariates that drive variability in drug response [171]. Traditional PopPK modelling relies on the direct application of parametric regression techniques, which often struggle to capture the complex relationships between covariates and drug pharmacokinetics [172,173]. The clear identification of most relevant covariates is essential for improving therapeutic drug monitoring and ensuring optimal treatment. Recent advancements in machine learning have introduced deep generative models, such as Variational Autoencoders (VAEs), which can learn latent representations of PK profiles while preserving meaningful structure in high-dimensional data [174,175]. In parallel, LASSO (Least Absolute Shrinkage and Selection) regression, a sparse linear model with L1 regularization [87,176], has been widely adopted for feature selection in biomedical applications due to its ability to eliminate non-contributory variables and enhance model interpretability [177]. In this study, we introduce a data-driven VAE-LASSO framework for covariate selection in PopPK modelling of tacrolimus, without requiring prior knowledge of the underlying pharmacokinetic model and its parameters. Tacrolimus is a commonly used immunosuppressant that plays a critical role in preventing graft rejection in patients who have undergone heart, liver or kidney transplantation [178]. Tacrolimus is a cornerstone immunosuppressive therapy in heart transplantation, merged as a valuable therapeutic alternative to ciclosporin following solid organ transplantation, where long-term graft survival relies on precise modulation of the recipient's immune response. Inadequate exposure increases the risk of acute cellular rejection, while excessive levels cause nephrotoxicity, neurotoxicity, metabolic disturbances, and heightened infection risk. Its narrow therapeutic index makes therapeutic drug monitoring essential, yet its pharmacokinetics are highly variable, influenced by demographic, genetic, and clinical factors. In heart transplant recipients, post-surgical changes, polypharmacy, and fluctuating organ function further complicate dosing, highlighting the need for robust modelling strategies to support personalized therapy. It presents a unique challenge for PK modelling due to its narrow therapeutic index and high inter-individual variability [179,180]. The VAE is used to encode PK signals into a latent space representation, while LASSO regression maps patient-specific covariates to the latent space, enabling the direct identification of key predictors of PK profiles. We systematically analyzed the impact of different regularization strengths (λ) on covariate selection

and assess the robustness of this approach in filtering out irrelevant ones. The findings of this study provide valuable insights into covariate selection by leveraging the latent representations learned through Variational Autoencoders. This approach enhances data-driven feature selection in PopPK modelling, with promising applications in precision dosing and personalized pharmacotherapy. Furthermore, we highlight the limitations of LASSO's linear assumption and discuss future directions for integrating latent space modelling with non-linear techniques to improve the accuracy and robustness of pharmacokinetic analyses.

4.2.3 Materials and Methods

Dataset Generation

1) Pharmacokinetics Modelling of Tacrolimus: The tacrolimus concentration at each time step, have been estimated using a one-compartment pharmacokinetic model depicted from (8. The generation of synthetic PK profiles was carried out in accordance with the settings outlined in the study by Chen et al. [179]. We generate Fast-Elimination PK profiles, with fixed dose D of 300 (mg) and elimination rate (k_e) calculated as $k_e = CL/V$, where V is the volume of distribution. The absorption rate constant k_a was fixed to 0.502 (h^{-1}) and the absorption time lag (tlag) was fixed to 0.346 (h). (9 is used to measure clearance CL for subject i at time j. In the equation, X_1 corresponds to *SNP* (Single Nucleotide Polymorphism) of the CYP3A5 genotypes. Following the encoding strategy adopted by Chen et al., the three genotypic variants (expressor, intermediate expressor, and non-expressor) are numerically coded as 1, 2, and 3, respectively. This numerical representation allows the inclusion of genetic polymorphism as a continuous covariate in the clearance model. X_2 refers to age as the subject age in year, X_3 is the albumin (alb) level in g/dL and X_4 the blood haemoglobin (hgb) concentration in g/dL of the subject i at time j. Volume has been estimated using (10.

$$C(t) = \frac{D}{V} \times \frac{k_a}{k_a - k_e} \times \left(e^{-k_e(t-t_D-t_{lag})} - e^{-k_a(t-t_D-t_{lag})}\right) \tag{8}$$

$$CL_{ij} = \theta_1 \times X_1^{\theta_2} \times X_2^{\theta_3} \times X_3^{\theta_4} \times X_4^{\theta_5} \times exp^{[nCL]} \tag{9}$$

$$V_{ij} = \theta_6 \times exp^{[nV]} \tag{10}$$

Where:

- $\theta_1 = 26.2$
- $X_1 = (SNP_{i,j})$ and $\theta_2 = 0.71$
- $X_2 = (\text{age}_{i,j} / 47)$ and $\theta_3 = -0.26$
- $X_3 = (\text{alb}_{i,j} / 4.1)$ and $\theta_4 = 0.35$
- $X_4 = (\text{hgb} / 125)$ and $\theta_5 = -0.29$
- $\theta_6 = 3726$

2) Covariates: The covariates used to generate the PK profiles include age, sex (male, female), weight, haemoglobin, albumin, CYP3A5 SNP (with three variants coded as 1,2,3), and ethnicity (race), categorized into Caucasian-American, African-American, Hispanic, Asian and other. Additionally, two extra random covariates (Called extra_1 and extra_2), generated as values uniformly distributed between 0 and 1, were introduced to further evaluate the system's behavior in covariate selection. Categorical covariates were sampled using a uniform random distribution to ensure equal probability among categories. Continuous variables were collected, sampling from a gaussian normal distribution, with each covariate assigned a specific mean (μ) and standard deviation (σ) (listed below).

- age: $\mu = 45.9$, $\sigma = 12.7$
- weight: $\mu = 82.9$, $\sigma = 20.8$
- haemoglobin: $\mu = 12.5$, $\sigma = 2.1$
- albumin: $\mu = 4.1$, $\sigma = 0.4$

3) Synthetic Data Generation: A total of 10,000 PK profiles, representing tacrolimus concentrations (mg/L) over a 48-hour period following dose administration, were created. Each with a specific set of covariates. In (9 and (10, $\exp^{[nCL]}$ and $\exp^{[nV]}$ represent the random effects of between-subject variability. These random effects were sampled from a normal distribution with a mean of 0 and standard deviations of 0.408 and 0.653, respectively. These pharmacokinetic signals were used to train the VAE, in order to construct a latent representation of all of them. A separate test set of 2,000 PK profiles was generated following the same procedure to assert the VAE model's PK profile reconstruction ability and generalization.

Model Architecture

The proposed framework consists of a Variational Autoencoder (VAE) trained to learn a latent representation of simulated pharmacokinetic (PK) profiles, followed by a Least Absolute Shrinkage and Selection Operator (LASSO) regression model that performs covariate selection, identifying those influencing the latent space reconstruction. The representation of the entire framework architecture can be seen in Figure 31.

1) PK Variational Autoencoder (VAE): Unlike a standard autoencoder, which compresses input data into a deterministic latent space, a VAE encodes the input into a probabilistic latent distribution. This probabilistic nature addresses the issue of non-regularized latent space of autoencoders and provides the generative capability to the entire space. The encoder maps PK profiles into a lower-dimensional latent space characterized by a mean (μ) and a standard deviation (σ), which parameterize a Gaussian distribution from which latent variables are sampled. The decoder reconstructs PK profiles from these latent variables, optimizing a combined loss function consisting of the reconstruction error and the Kullback-Leibler (KL) divergence, which regularizes the latent space distribution [181].

2) Latent Space Regression: The LASSO regression model, determine the influence on each covariate in the latent space reconstruction by adding a regularization penality term the L1 regularization. The penality term forces regression coefficients to become exactly zero, performing data-driven feature selection [182]. In the (11 we can see the classical ordinary least squares (OLS) function used to estimate the coefficient vector $\hat{\beta}$ that represents the weights assigned to each predictor (i.e. each covariate). $\lambda \sum_{j=1}^{p} |\beta_j|$ is the L1 regularization penalty term, that controls the sparsity of the model. λis the crucial parameter that determine the strength of the penalty term.

$$\hat{\beta} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} X_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \quad (11)$$

Experimental procedure

1) Covariate Preprocessing: Prior to model training, categorical covariates were processed using one-hot encoding, while continuous covariates were normalized using min-max scaling to ensure consistency across features.

2) VAE Model Training: The VAE model was trained using the 10,000 PK profiles as input, to obtain a representative latent space. The training process was optimized using mean absolute error (MAE) as reconstruction loss to minimize differences between input and output PK signals and a KL-divergence term to enforce smoothness in the latent space distribution. We evaluated VAE architecture on the generated test set of 2,000 tacrolimus PK profiles. The model achieves an optimal mean absolute percentage error (MAPE) of 2.26% (measured metrics are collected in Table I) in PK-profiles reconstruction. Metrics have been measured using Python's scikit-learn library. Figure 32 shows the comparison of the test set (panel A) and the corresponding reconstructed profiles generated by the VAE (panel B). Each line represents one of the 2,000 individual PK profiles.

3) LASSO Regression model: The LASSO regression model was developed using Python's scikit-learn library [183]. It was applied to establish a direct mapping between patient-specific covariates and the latent space representations of the 10,000 PK-tacrolimus profiles used as training set for the VAE. To evaluate the impact of regularization strength on covariate selection, different values of the λ parameter were tested. Lower λ values allowed for a larger number of covariates to be retained, whereas higher λ values forced greater sparsity, removing less relevant features. For each tested λ value, the weights assigned to each covariate were collected and analyzed. We systematically explored a suitable range of regularization strengths and conducted an in-depth analysis of covariate selection for the following λ values: 0.0001,0.002,0.005,0.008,0.01,0.1,1.0.

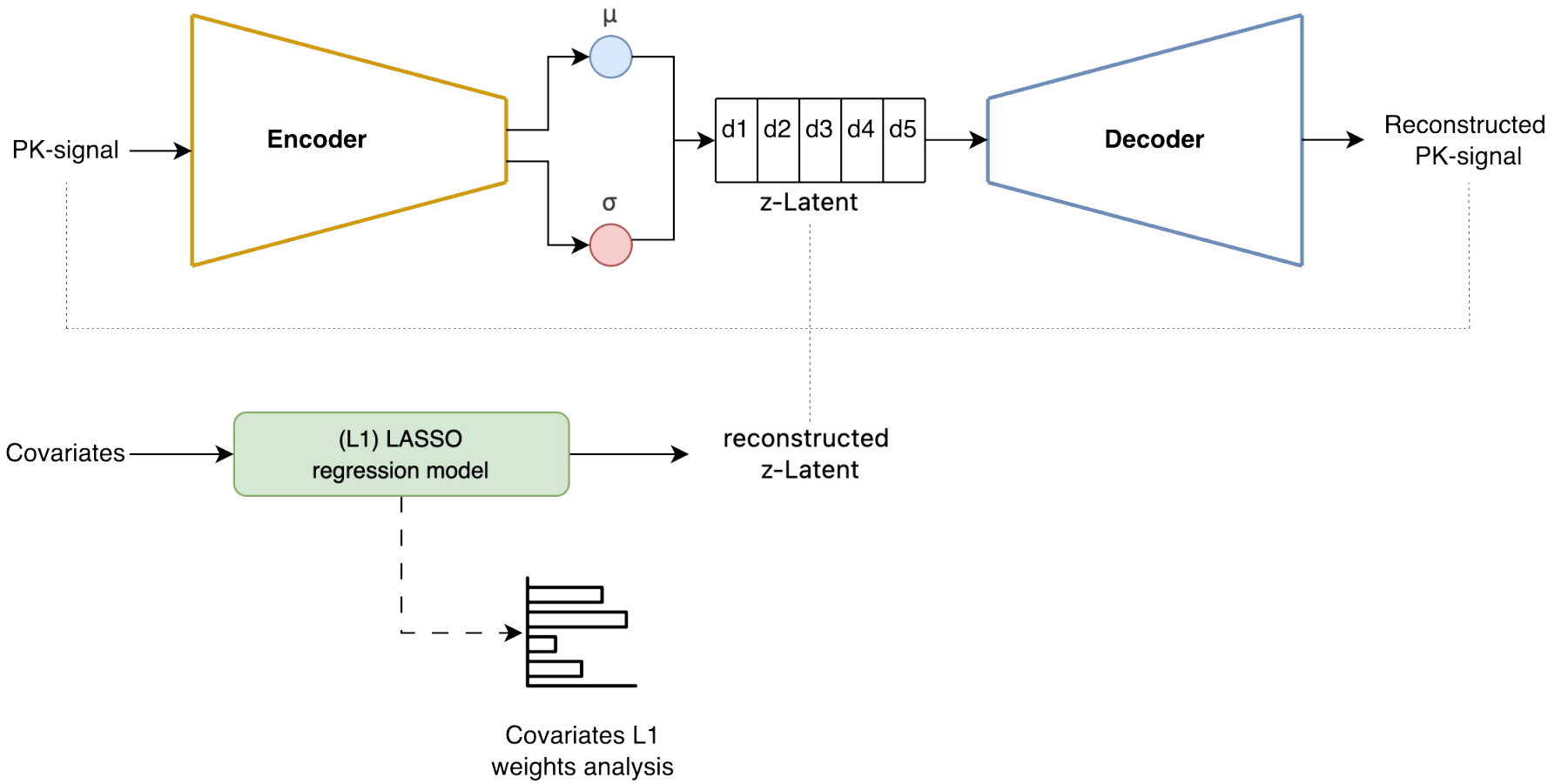


Figure 31. Graphical representation of the VAE-LASSO framework. The Variational Autoencoder consists of an encoder (left) that compresses pharmacokinetic signals profiles into a latent representation parameterized by a mean (μ) and standard deviation (σ). The latent variables are then sampled and passed tothe decoder (right), which reconstructs the original PK profiles. A LASSO regression model (bottom) is trained to predict the latent representation using patient-specific covariates, allowing for covariate selection by shrinking irrelevant features to zero.

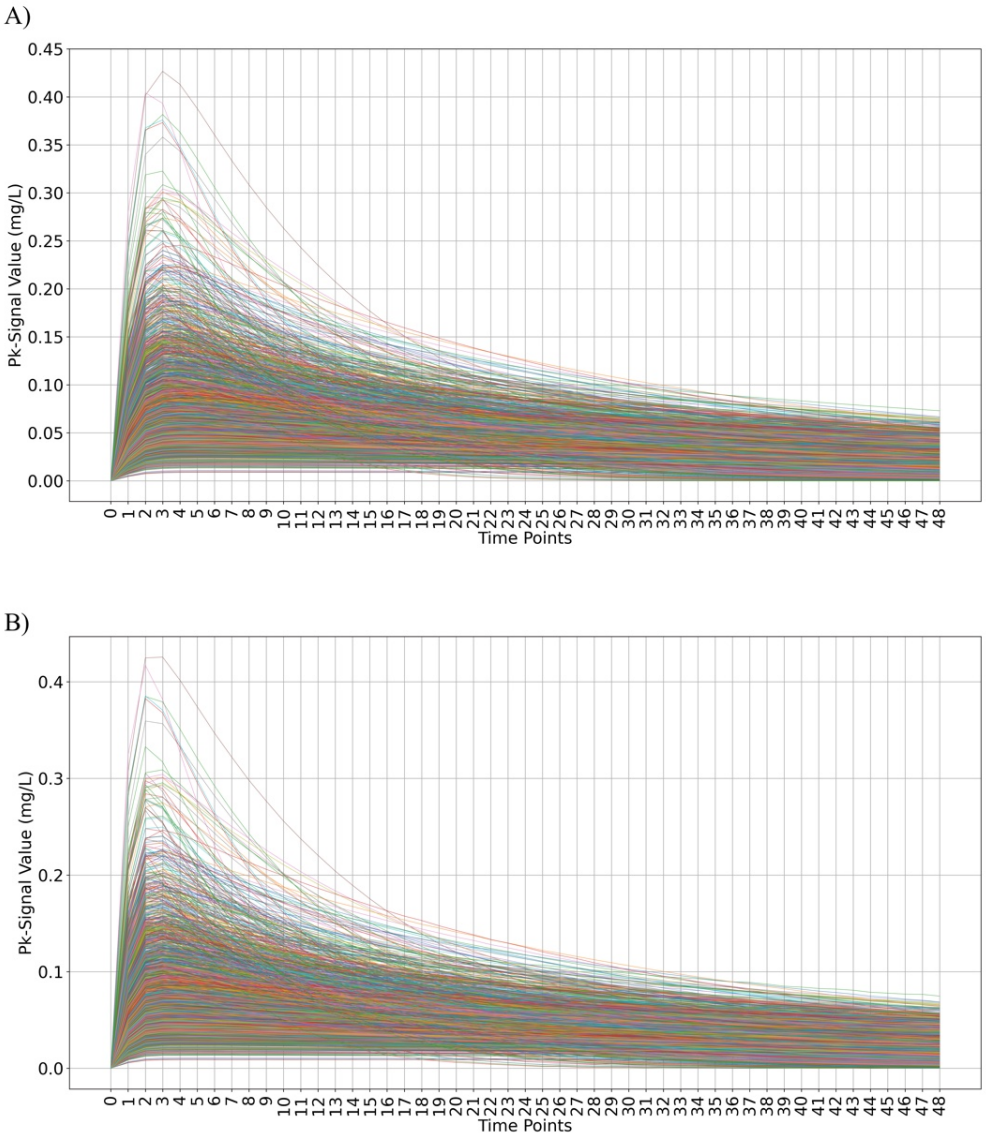
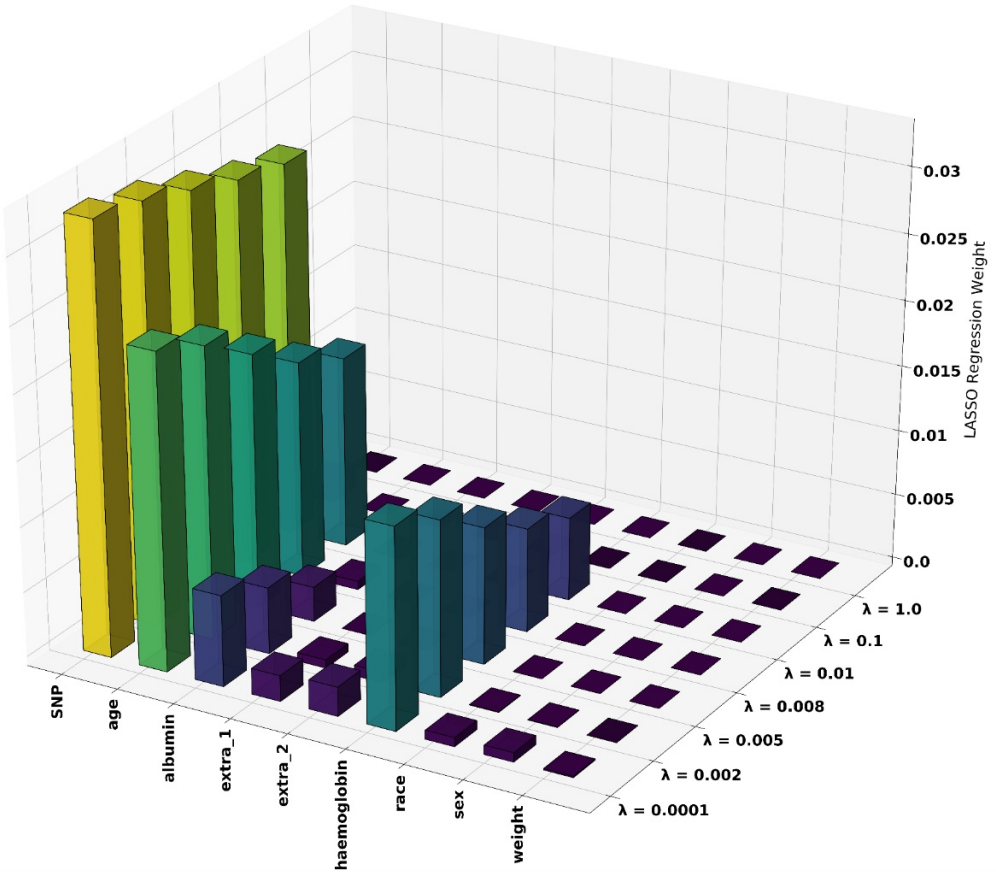


Figure 32. (A) Original tacrolimus PK profiles generated from the simulated dataset. (B) Reconstructed PK profiles obtained from the Variational Autoencoder (VAE). The y-axis represents tacrolimus concentration in mg/L over time. The optimal reconstruction that closely follow the original data, demonstrates the VAE's ability to capture the underlying PK dynamics.

Metric	VAE Model
MAE (Mean Absolute Error)	0.0009 (mg/L)
MAPE (Mean Absolute Percentage Error)	2.26 %

Table 9. PK profiles reconstruction error from VAE model



λ	SNP	Race	Sex	Extra_1	Extra_2	Albumin	Haemoglobin	Weight	Age
0.0001	✔	✔	✔	✔	✔	✔	✔	✔	✔
0.002	✔			✔	✔	✔	✔		✔
0.005	✔					✔	✔		✔
0.008	✔					✔	✔		✔
0.01	✔						✔		✔
0.1	✔								
1.0									

Figure 33. 3D Visualization of LASSO Regression weights across covariates and lambda values. The height and color intensity of the bars represent the importance of each covariate at different levels of regularization λ. Brighter and taller bars indicate covariates with stronger influence (i.e., higher regression weights), while darker, shorter bars correspond to covariates with little or no contribution.The x-axis lists the covariates, y-axis the different values of the regularization parameter λand the z-axis the LASSO regression weights. The table denote with markers the covariates retained at each regularization level.

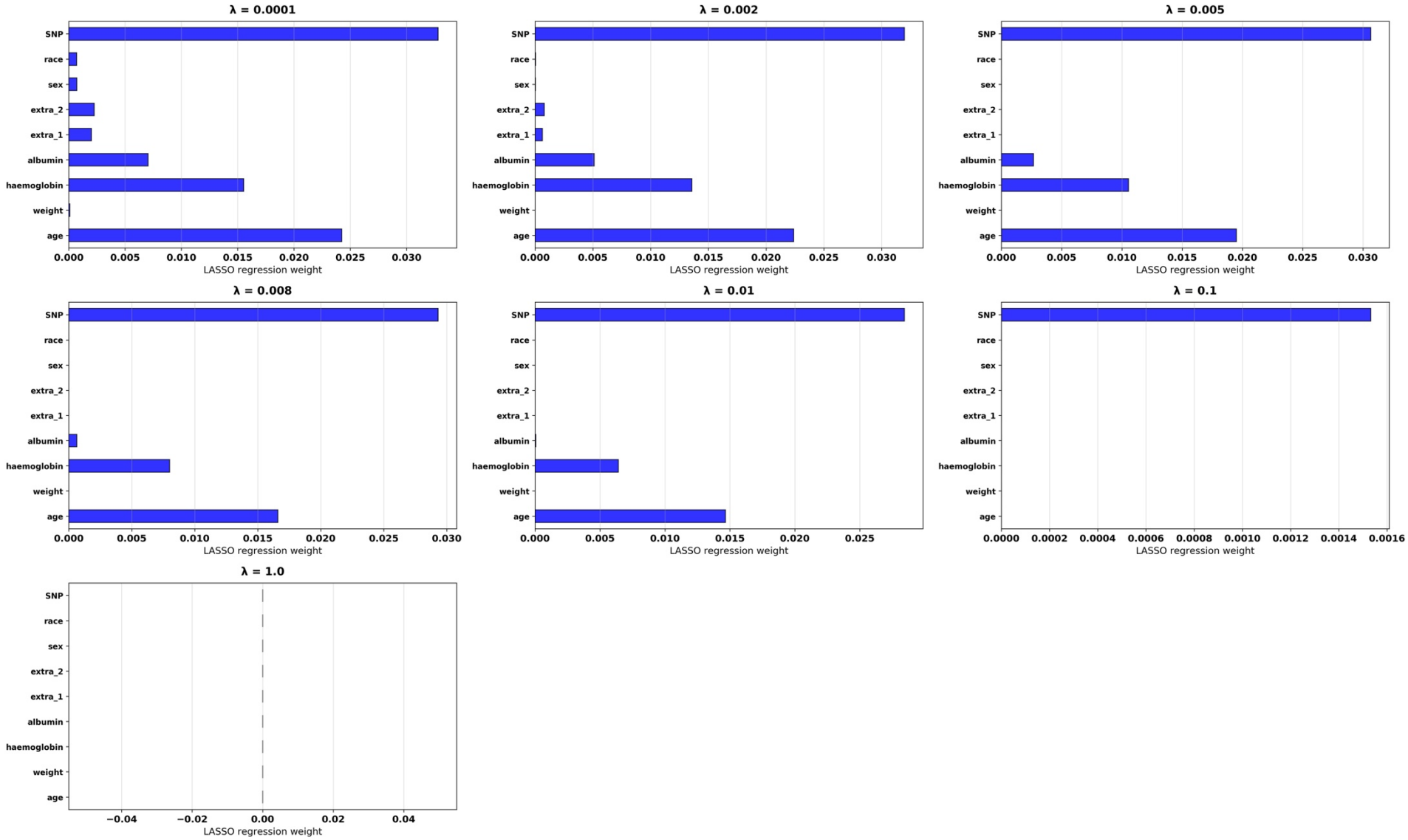


Figure 34. 2D Visualization of LASSO Regression weights across covariates and regularization values (λ). Each subplot corresponds to a different value of the regularization parameter λ, showing the LASSO regression weights assigned to each covariate. The x-axis represents the LASSO regression weight, while the y-axis lists the covariates. As λ increases, more covariates are progressively shrunk toward zero, indicating reduced importance.

4.2.4 Results

Our results demonstrate that the LASSO regression model effectively identifies key covariates in (9 that characterize tacrolimus pharmacokinetic profiles by reconstructing the latent space generated by the VAE. Figure 33 and Figure 34 illustrate the covariates selected by the model and the effect of varying the regularization parameter λ. As λ increases, the model progressively eliminates weaker covariates and the two randomly generated noise variables (extra 1 and extra 2), confirming the robustness of LASSO in filtering out non-contributory features. The SNP, age, hemoglobin, and albumin covariates are consistently retained across different λ values, emphasizing their fundamental role in tacrolimus pharmacokinetics. The optimal λ value is determined by balancing model interpretability with feature selection efficiency. Lower λ values allow more covariates to contribute, potentially capturing minor effects, whereas higher λ values enforce greater sparsity, retaining only the most influential predictors. To further assess the effectiveness of LASSO-based covariate selection, we utilized the reconstructed latent space to decode PK profiles via the VAE decoder. However, the reconstructed PK profiles exhibited insufficient accuracy, primarily due to the linear nature of LASSO regression. Since pharmacokinetic signals are governed by complex, non-linear dynamics, the simplified linear mapping imposed by LASSO does not allow the complete correct reconstruction of the PK profiles. Despite this limitation, our findings indicate that the LASSO-reconstructed latent space provides an efficient structured and interpretable representation that enables efficient covariate selection. Importantly, the selected covariates also align with established clinical determinants of tacrolimus exposure in the cardiac transplantation setting. Age emerged as a relevant factor, with differences in drug clearance observed in both pediatric and elderly patients. Albumin levels were consistently retained, reflecting their role in modulating the unbound fraction of tacrolimus. Hemoglobin was also identified, representing a surrogate for hematocrit and influencing tacrolimus whole-blood distribution, particularly in transplant patients who frequently experience anemia or transfusions.

4.2.3 Discussion

Our findings demonstrate that the proposed VAE-LASSO framework effectively integrates latent space modelling with sparse regression techniques for covariate selection in population pharmacokinetics (PopPK). Unlike traditional PopPK approaches, which rely heavily on parametric regression models with predefined structural assumptions, our method is model-free and data-driven, requiring no prior knowledge of the underlying pharmacokinetic equations or covariate interactions. This distinction is crucial, as traditional PopPK models often rely on stepwise covariate selection procedures such as forward selection or backward elimination which, despite their simplicity, are prone to overfitting and may fail to capture complex, nonlinear relationships in high-dimensional data [184,185]. In contrast, our approach leverages a Variational Autoencoder to learn a compact and structured latent representation directly from the pharmacokinetic profiles, enabling the discovery of hidden patterns without requiring prior assumptions about the underlying model structure. The VAE-LASSO framework identified clinically relevant covariates, which are

consistent with established evidence in cardiac transplantation. Age has long been recognized as a determinant of tacrolimus pharmacokinetics, with both pediatric and elderly patients exhibiting altered clearance rates due to developmental or degenerative changes in organ and metabolic function. This variability is particularly relevant in heart transplantation, where the patient population is heterogeneous and dosing strategies must account for these physiological differences. Tacrolimus is highly bound to plasma proteins, and fluctuations in albumin directly affect the proportion of free, pharmacologically active drug Hypoalbuminemia is frequently observed in heart transplant recipients during the early post-surgical period, which may contribute to increased exposure to the unbound fraction of tacrolimus and a heightened risk of toxicity [186]. The consistent selection of hemoglobin by our model highlights its relevance as a practical and easily measurable clinical marker that can inform dose adjustment. Hemoglobin reflects hematocrit and plays a crucial role in tacrolimus whole-blood distribution. Low hemoglobin or anemia, which is common in heart transplant patients due to surgical blood loss, transfusions, and concomitant therapies, has been associated with changes in drug disposition and intra-patient variability [187]. Perhaps most importantly, the genetic polymorphism covariates retained by the LASSO model are supported by a large body of evidence demonstrating the impact of CYP3A5 genotype on tacrolimus metabolism. Patients carrying the CYP3A5 allele (“expressors”) require significantly higher doses to achieve target trough concentrations compared with non-expressors, a finding consistently reported across solid organ transplantation and confirmed in adult heart transplant recipients [188] [189].Taken together, these findings show that the covariates highlighted by our framework are not only statistically robust but also biologically and clinically meaningful. Despite its effectiveness in feature selection, the linear nature of LASSO regression limits its ability to fully reconstruct the latent space and, consequently, the pharmacokinetic profiles. This limitation arises from the non-linearity inherent in PK systems, which LASSO fails to capture. To address this limitation, we are currently exploring an alternative architecture based on a Conditional Variational Autoencoder (CVAE), in which a LASSO-regularized layer is incorporated to condition the reconstruction of the latent space generated by the encoder. This design aims to more effectively capture the nonlinear dynamics of pharmacokinetic time concentration profiles while maintaining covariate interpretability. In parallel, we are evaluating the architecture on a more complex compartmental PK model to further assess its applicability and robustness. Nevertheless, the model provides a structured and interpretable latent representation that simplifies covariate selection, making it a valuable tool for clinical decision-making. It also highlights the potential of latent space representations in uncovering underlying pharmacokinetic patterns, facilitating a data-driven approach to identifying key patient-specific factors influencing drug metabolism. By leveraging the outcome of this study, future works will aim to further explore the integration of latent space modelling with other multiple non-linear based techniques and improving the interpretability of latent space representations for enhance therapeutic decision- making. Overall, this study demonstrates the potential of integrating deep learning and sparse regression to enhance covariate selection in pharmacokinetic modelling, paving the way for more robust and interpretable machine learning approaches in clinical pharmacology.

5 Conclusions

This thesis has investigated the development and application of machine learning models for the prediction of major adverse cardiac events in heart transplant patients, with particular focus on designing approaches that remain robust in data-scarce scenarios. A central contribution is the implementation of a classification framework that integrates data augmentation and feature selection to improve the trade-off between model accuracy and interpretability. This model, validated through a comprehensive bootstrap analysis on a publicly available dataset, laid the foundation for the AugPred pipeline, a tool designed on the same concept to support post-transplant surveillance by combining synthetic data generation and feature selection for robust classification with bioinformatic analysis of selected features. Building on this methodology, the thesis introduced a multilayer network model to characterize the topological structure and regulatory dynamics between microRNAs and their target mRNAs, integrating multiple data sources from public databases and real measured molecular level concentration profiles, revealing phenotype-specific rewiring in rejection and infection cases. Furthermore, a deep generative architecture called VAE-LASSO was developed to analyze pharmacokinetic time-concentration profiles of tacrolimus immunosuppressant. This model demonstrated the potential of latent space representations for identifying clinically relevant covariates, opening the way to more personalized and interpretable immunosuppressive therapy.
The results obtained across Chapters 2 to 4 can be summarized as follows:

1. The developed architecture and subsequent bootstrap analysis demonstrated the effective application of data augmentation and feature selection to balance the trade-off between classification accuracy and interpretability when working with high-dimensional datasets. The systematic evaluation across multiple binary classification tasks confirmed the discriminative power of selecting a restricted number of features, enabling robust classification performance while preserving biological interpretability across different pathological conditions.

2. The AugPred pipeline demonstrated strong applicability in the heart transplant setting, even under conditions of limited data availability. By leveraging synthetic data augmentation and feature selection, the model effectively discriminated between acute cellular rejection and cytomegalovirus infection using transcriptomic profiles derived from microarray data. Notably, the selected microRNAs were not only predictive but also biologically meaningful, as confirmed by downstream pathway enrichment analyses. The pipeline maintained high classification performance across varying sample sizes and proved adaptable to different clinical and experimental scenarios. These findings support its potential integration into post-transplant monitoring workflows and underscore its ability to identify mechanistically relevant biomarkers in cardiac transplantation.

3. The multilayer network model developed, representing miRNA–mRNA interactions offered a better global system-level view of post-transcriptional molecular regulatory architectures in heart transplanted patients. By modeling miRNAs and mRNAs as nodes of intra and interconnected layers, and applying PageRank centrality within this multilayer framework,

the study identified phenotype-specific shifts in node centrality modulated from a subset of microRNAs, pointing to regulatory reorganization with weakened post-transcriptional control in acute cellular rejection cases. This network-based approach revealed the non-redundant and dynamic role of microRNAs as modulators of gene expression across clinical phenotypes, extending beyond conventional differential expression analysis.

4. The VAE-LASSO framework demonstrated the potential of combining deep generative learning models with interpretable statistical learning to uncover clinically relevant covariates from high-dimensional data. By using variational autoencoder (VAE) to model the latent space of pharmacokinetic (PK) time concentration profiles, followed by LASSO regression to identify informative covariates, this approach enabled the extraction of covariates influencing PK variability in tacrolimus immunosuppressant absorption. The method effectively bridged generative representation learning with sparse feature selection, providing a scalable strategy for more personalized dosing.

Taken together, these findings highlight the potential of machine learning models to provide clinically meaningful support systems, even in settings constrained by limited sample size. It shows the valuable implication of the complex network theory in transplant research, by pointing out the effectivenes to model trascriptomic interaction to investigate the topology and dynamic of the complex interconnect biological system that declare the different post-transplant outcomes. The main contribution of this thesis lies in demonstrating the feasibility and effectiveness of combining computational methods such as machine learning, deep learning and network theory, with clinical application, to enhance the support in the decision making. This perspective not only advances methodological innovation but also provides a translational method, showing how computational tools can support clinicians in the management of heart transplant patients. Nevertheless, some limitations must be acknowledged. Among the main challenges are the small cohort sizes which determine the goodness of the data and the complexity of acquiring biologically relevant information from clinical procedures. Additionally, the intrinsic interconnectedness and dynamism of biological systems make them particularly complex to model. These limitations also suggest clear future directions. Explore the application of multimodal learning frameworks capable of integrating heterogeneous data sources, such as clinical variables, omics data (genomics, transcriptomics, proteomics, etc.), medical images, and longitudinal measurements, while also extracting informative features related to the topology and dynamics of interactions across different system levels through complex network models. such integration is fundamental to capturing the structure and dynamic nature of biological systems in the context of transplantation, ultimately supporting more personalized and context-aware diagnostic strategies. In conclusion, this doctoral thesis contributes novel tools, methodological frameworks, and validation strategies that may support the development of precision medicine approaches in transplantation, enabling better patient stratification, earlier detection of adverse events, and ultimately, more personalized and effective care. By enhancing the capacity to anticipate adverse cardiac events and uncover latent regulatory mechanisms, these approaches contribute to a deeper understanding of post-transplant trajectories and open the way to more personalized, accurate, and effective patient care.

References

[1] Reich BH, Wee SY. Searching for Knowledge in the Pmbok® Guide. Project Management Journal 2006;37:11–26. https://doi.org/10.1177/875697280603700203.

[2] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017;2:230–43. https://doi.org/10.1136/svn-2017-000101.

[3] Ahmed Z, Mohamed K, Zeeshan S, Dong X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database 2020;2020:baaa010. https://doi.org/10.1093/database/baaa010.

[4] Peiffer-Smadja N, Rawson TM, Ahmad R, Buchard A, Georgiou P, Lescure F-X, et al. Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. Clinical Microbiology and Infection 2020;26:584–95. https://doi.org/10.1016/j.cmi.2019.09.009.

[5] Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Briefings in Bioinformatics 2018;19:1236–46. https://doi.org/10.1093/bib/bbx044.

[6] Riedl JM, Moik F, Esterl T, Kostmann SM, Gerger A, Jost PJ. Molecular diagnostics tailoring personalized cancer therapy-an oncologist's view. Virchows Arch 2024;484:169–79. https://doi.org/10.1007/s00428-023-03702-7.

[7] Loupy A, Duong Van Huyen JP, Hidalgo L, Reeve J, Racapé M, Aubert O, et al. Gene Expression Profiling for the Identification and Classification of Antibody-Mediated Heart Rejection. Circulation 2017;135:917–35. https://doi.org/10.1161/CIRCULATIONAHA.116.022907.

[8] Lipkova J, Chen TY, Lu MY, Chen RJ, Shady M, Williams M, et al. Deep learning-enabled assessment of cardiac allograft rejection from endomyocardial biopsies. Nat Med 2022;28:575–82. https://doi.org/10.1038/s41591-022-01709-2.

[9] Lisboa PJG, Jayabalan M, Ortega-Martorell S, Olier I, Medved D, Nilsson J. Enhanced survival prediction using explainable artificial intelligence in heart transplantation. Sci Rep 2022;12:19525. https://doi.org/10.1038/s41598-022-23817-2.

[10] Mohammadi I, Farahani S, Karimi A, Jahanian S, Firouzabadi SR, Alinejadfard M, et al. Mortality prediction of heart transplantation using machine learning models: a systematic review and meta-analysis. Front Artif Intell 2025;8:1551959. https://doi.org/10.3389/frai.2025.1551959.

[11] Khush KK, Hsich E, Potena L, Cherikh WS, Chambers DC, Harhay MO, et al. The International Thoracic Organ Transplant Registry of the International Society for Heart and Lung Transplantation: Thirty-eighth adult heart transplantation report — 2021; Focus on recipient characteristics. The Journal of Heart and Lung Transplantation 2021;40:1035–49. https://doi.org/10.1016/j.healun.2021.07.015.

[12] Lund LH, Khush KK, Cherikh WS, Goldfarb S, Kucheryavaya AY, Levvey BJ, et al. The Registry of the International Society for Heart and Lung Transplantation: Thirty-fourth Adult Heart Transplantation Report—2017; Focus Theme: Allograft ischemic time. The Journal of Heart and Lung Transplantation 2017;36:1037–46. https://doi.org/10.1016/j.healun.2017.07.019.

[13] Peled Y, Ducharme A, Kittleson M, Bansal N, Stehlik J, Amdani S, et al. International Society for Heart and Lung Transplantation Guidelines for the Evaluation and Care of Cardiac Transplant Candidates—2024. The Journal of Heart and Lung Transplantation 2024;43:1529-1628.e54. https://doi.org/10.1016/j.healun.2024.05.010.

[14] Khush KK, Cherikh WS, Chambers DC, Harhay MO, Hayes D, Hsich E, et al. The International Thoracic Organ Transplant Registry of the International Society for Heart and Lung Transplantation: Thirty-sixth adult heart transplantation report — 2019; focus theme: Donor and recipient size match. The Journal of Heart and Lung Transplantation 2019;38:1056–66. https://doi.org/10.1016/j.healun.2019.08.004.

[15] Khush KK, Cherikh WS, Chambers DC, Goldfarb S, Hayes D, Kucheryavaya AY, et al. The International Thoracic Organ Transplant Registry of the International Society for Heart and Lung Transplantation: Thirty-fifth Adult Heart Transplantation Report—2018; Focus Theme: Multiorgan Transplantation. The Journal of Heart and Lung Transplantation 2018;37:1155–68. https://doi.org/10.1016/j.healun.2018.07.022.

[16] Medved D, Ohlsson M, Höglund P, Andersson B, Nugues P, Nilsson J. Improving prediction of heart transplantation outcome using deep learning techniques. Sci Rep 2018;8:3613. https://doi.org/10.1038/s41598-018-21417-7.

[17] Oztekin A, Delen D, Kong Z (James). Predicting the graft survival for heart–lung transplantation patients: An integrated data mining methodology. International Journal of Medical Informatics 2009;78:e84–96. https://doi.org/10.1016/j.ijmedinf.2009.04.007.

[18] Stewart S, Winters GL, Fishbein MC, Tazelaar HD, Kobashigawa J, Abrams J, et al. Revision of the 1990 Working Formulation for the Standardization of Nomenclature in the Diagnosis of Heart Rejection. The Journal of Heart and Lung Transplantation 2005;24:1710–20. https://doi.org/10.1016/j.healun.2005.03.019.

[19] Costanzo MR, Costanzo MR, Dipchand A, Starling R, Anderson A, Chan M, et al. The International Society of Heart and Lung Transplantation Guidelines for the care of heart transplant recipients. The Journal of Heart and Lung Transplantation 2010;29:914–56. https://doi.org/10.1016/j.healun.2010.05.034.

[20] Berry GJ, Burke MM, Andersen C, Bruneval P, Fedrigo M, Fishbein MC, et al. The 2013 International Society for Heart and Lung Transplantation Working Formulation for the standardization of nomenclature in the pathologic diagnosis of antibody-mediated rejection in heart transplantation. The Journal of Heart and Lung Transplantation 2013;32:1147–62. https://doi.org/10.1016/j.healun.2013.08.011.

[21] Mi H, Sivagnanam S, Ho WJ, Zhang S, Bergman D, Deshpande A, et al. Computational methods and biomarker discovery strategies for spatial proteomics: a review in immuno-oncology. Briefings in Bioinformatics 2024;25:bbae421. https://doi.org/10.1093/bib/bbae421.

[22] Holzhauser L, DeFilippis EM, Nikolova A, Byku M, Contreras JP, De Marco T, et al. The End of Endomyocardial Biopsy? JACC: Heart Failure 2023;11:263–76. https://doi.org/10.1016/j.jchf.2022.11.002.

[23] Shanes JG, Ghali J, Billingham ME, Ferrans VJ, Fenoglio JJ, Edwards WD, et al. Interobserver variability in the pathologic interpretation of endomyocardial biopsy results. Circulation 1987;75:401–5. https://doi.org/10.1161/01.CIR.75.2.401.

[24] Marino J, Paster J, Benichou G. Allorecognition by T Lymphocytes and Allograft Rejection. Front Immunol 2016;7. https://doi.org/10.3389/fimmu.2016.00582.

[25] Stewart S, Winters GL, Fishbein MC, Tazelaar HD, Kobashigawa J, Abrams J, et al. Revision of the 1990 Working Formulation for the Standardization of Nomenclature in the Diagnosis of

Heart Rejection. The Journal of Heart and Lung Transplantation 2005;24:1710–20. https://doi.org/10.1016/j.healun.2005.03.019.

[26] Potena L, Solidoro P, Patrucco F, Borgese L. Treatment and prevention of cytomegalovirus infection in heart and lung transplantation: an update. Expert Opinion on Pharmacotherapy 2016;17:1611–22. https://doi.org/10.1080/14656566.2016.1199684.

[27] Mocarski ES. Immunomodulation by cytomegaloviruses: manipulative strategies beyond evasion. Trends in Microbiology 2002;10:332–9. https://doi.org/10.1016/S0966-842X(02)02393-4.

[28] Razonable RR, Humar A. Cytomegalovirus in solid organ transplant recipients—Guidelines of the American Society of Transplantation Infectious Diseases Community of Practice. Clinical Transplantation 2019;33:e13512. https://doi.org/10.1111/ctr.13512.

[29] Kotton CN, Kumar D, Caliendo AM, Huprikar S, Chou S, Danziger-Isakov L, et al. The Third International Consensus Guidelines on the Management of Cytomegalovirus in Solid-organ Transplantation. Transplantation 2018;102:900–31. https://doi.org/10.1097/TP.0000000000002191.

[30] Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 1993;75:843–54. https://doi.org/10.1016/0092-8674(93)90529-Y.

[31] Gargiulo P, Marzano F, Salvatore M, Basile C, Buonocore D, Parlati ALM, et al. MicroRNAs: diagnostic, prognostic and therapeutic role in heart failure—a review. ESC Heart Failure 2023;10:753–61. https://doi.org/10.1002/ehf2.14153.

[32] Wojciechowska A, Osiak A, Kozar-Kamińska K. MicroRNA in cardiovascular biology and disease. Adv Clin Exp Med 2017;26:868–74. https://doi.org/10.17219/acem/62915.

[33] Gargiulo P, Marzano F, Salvatore M, Basile C, Buonocore D, Parlati ALM, et al. MicroRNAs: diagnostic, prognostic and therapeutic role in heart failure—a review. ESC Heart Failure 2023;10:753–61. https://doi.org/10.1002/ehf2.14153.

[34] Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 2014;15:509–24. https://doi.org/10.1038/nrm3838.

[35] Fung EC, Butt AN, Eastwood J, Swaminathan R, Sodi R. Circulating microRNA in cardiovascular disease. Advances in Clinical Chemistry, vol. 91, Elsevier; 2019, p. 99–122. https://doi.org/10.1016/bs.acc.2019.03.003.

[36] Nováková T, Macháčková T, Novák J, Hude P, Godava J, Žampachová V, et al. Identification of a Diagnostic Set of Endomyocardial Biopsy microRNAs for Acute Cellular Rejection Diagnostics in Patients after Heart Transplantation Using Next-Generation Sequencing. Cells 2019;8:1400. https://doi.org/10.3390/cells8111400.

[37] Di Francesco A, Fedrigo M, Santovito D, Natarelli L, Castellani C, De Pascale F, et al. MicroRNA signatures in cardiac biopsies and detection of allograft rejection. The Journal of Heart and Lung Transplantation 2018;37:1329–40. https://doi.org/10.1016/j.healun.2018.06.010.

[38] Duong Van Huyen J-P, Tible M, Gay A, Guillemain R, Aubert O, Varnous S, et al. MicroRNAs as non-invasive biomarkers of heart transplant rejection. European Heart Journal 2014;35:3194–202. https://doi.org/10.1093/eurheartj/ehu346.

[39] Kennel PJ, Yahi A, Naka Y, Mancini DM, Marboe CC, Max K, et al. Longitudinal profiling of circulating miRNA during cardiac allograft rejection: a proof-of-concept study. ESC Heart Failure 2021;8:1840–9. https://doi.org/10.1002/ehf2.13238.

[40] Sukma Dewi I, Hollander Z, Lam KK, McManus J-W, Tebbutt SJ, Ng RT, et al. Association of Serum MiR-142-3p and MiR-101-3p Levels with Acute Cellular Rejection after Heart Transplantation. PLoS ONE 2017;12:e0170842. https://doi.org/10.1371/journal.pone.0170842.

[41] Sukma Dewi I, Celik S, Karlsson A, Hollander Z, Lam K, McManus J-W, et al. Exosomal miR-142-3p is increased during cardiac allograft rejection and augments vascular permeability through down-regulation of endothelial RAB11FIP2 expression. Cardiovasc Res 2017:cvw244. https://doi.org/10.1093/cvr/cvw244.

[42] Constanso-Conde I, Hermida-Prieto M, Barge-Caballero E, Núñez L, Pombo-Otero J, Suárez-Fuentetaja N, et al. Circulating miR-181a-5p as a new biomarker for acute cellular rejection in heart transplantation. The Journal of Heart and Lung Transplantation 2020;39:1100–8. https://doi.org/10.1016/j.healun.2020.05.018.

[43] Connor KL, O'Sullivan ED, Marson LP, Wigmore SJ, Harrison EM. The Future Role of Machine Learning in Clinical Transplantation. Transplantation 2021;105:723–35. https://doi.org/10.1097/TP.0000000000003424.

[44] Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electron Markets 2021;31:685–95. https://doi.org/10.1007/s12525-021-00475-2.

[45] Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006.

[46] Murphy KP. Machine Learning - A Probabilistic Perspective. Cambridge: MIT Press; 2014.

[47] Zhu XJ. Semi-supervised learning literature survey 2005.

[48] Sutton RS, Barto A. Reinforcement learning: an introduction. Second edition. Cambridge, Massachusetts London, England: The MIT Press; 2020.

[49] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44. https://doi.org/10.1038/nature14539.

[50] Hornik K. Approximation capabilities of multilayer feedforward networks. Neural Networks 1991;4:251–7. https://doi.org/10.1016/0893-6080(91)90009-T.

[51] Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med 2019;25:24–9. https://doi.org/10.1038/s41591-018-0316-z.

[52] Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. IEEE J Biomed Health Inform 2018;22:1589–604. https://doi.org/10.1109/JBHI.2017.2767063.

[53] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention Is All You Need 2017. https://doi.org/10.48550/ARXIV.1706.03762.

[54] Ghonge NP, Goyal N, Vohra S, Chowdhury V. Renal transplant evaluation: multimodality imaging of post-transplant complications. BJR 2021;94:20201253. https://doi.org/10.1259/bjr.20201253.

[55] Liou L, Mostofsky E, Lehman L, Salia S, Barrera FJ, Wei Y, et al. Survival machine learning methods for mortality prediction after heart transplantation in the contemporary era. PLoS ONE 2025;20:e0313600. https://doi.org/10.1371/journal.pone.0313600.

[56] Lipkova J, Chen TY, Lu MY, Chen RJ, Shady M, Williams M, et al. Deep learning-enabled assessment of cardiac allograft rejection from endomyocardial biopsies. Nat Med 2022;28:575–82. https://doi.org/10.1038/s41591-022-01709-2.

[57] Seraphin TP, Luedde M, Roderburg C, Van Treeck M, Scheider P, Buelow RD, et al. Prediction of heart transplant rejection from routine pathology slides with self-supervised deep

learning. European Heart Journal - Digital Health 2023;4:265–74. https://doi.org/10.1093/ehjdh/ztad016.

[58] Yang H, Yang M, Chen J, Yao G, Zou Q, Jia L. Multimodal deep learning approaches for precision oncology: a comprehensive review. Briefings in Bioinformatics 2024;26:bbae699. https://doi.org/10.1093/bib/bbae699.

[59] Di Martino F, Delmastro F. Explainable AI for clinical and remote health applications: a survey on tabular and time series data. Artif Intell Rev 2023;56:5261–315. https://doi.org/10.1007/s10462-022-10304-3.

[60] El Naqa I. Prospective clinical deployment of machine learning in radiation oncology. Nat Rev Clin Oncol 2021;18:605–6. https://doi.org/10.1038/s41571-021-00541-w.

[61] Barabási A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–13. https://doi.org/10.1038/nrg1272.

[62] Hu JX, Thomas CE, Brunak S. Network biology concepts in complex disease comorbidities. Nat Rev Genet 2016;17:615–29. https://doi.org/10.1038/nrg.2016.87.

[63] De Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter MA, et al. Mathematical Formulation of Multilayer Networks. Phys Rev X 2013;3:041022. https://doi.org/10.1103/PhysRevX.3.041022.

[64] Barabási A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–13. https://doi.org/10.1038/nrg1272.

[65] Page L, Brin S, Motwani R, Winograd T. The PageRank Citation Ranking: Bringing Order to the Web. Stanford InfoLab; 1999.

[66] Li J, Zhao PX. Mining Functional Modules in Heterogeneous Biological Networks Using Multiplex PageRank Approach. Front Plant Sci 2016;7. https://doi.org/10.3389/fpls.2016.00903.

[67] Barthélemy M. Betweenness centrality in large complex networks. Eur Phys J B 2004;38:163–8. https://doi.org/10.1140/epjb/e2004-00111-4.

[68] De Domenico M, Solé-Ribalta A, Omodei E, Gómez S, Arenas A. Ranking in interconnected multilayer networks reveals versatile nodes. Nat Commun 2015;6:6868. https://doi.org/10.1038/ncomms7868.

[69] Duch J, Arenas A. Community detection in complex networks using extremal optimization. Phys Rev E 2005;72:027104. https://doi.org/10.1103/PhysRevE.72.027104.

[70] Artime O, Grassia M, De Domenico M, Gleeson JP, Makse HA, Mangioni G, et al. Robustness and resilience of complex networks. Nat Rev Phys 2024;6:114–31. https://doi.org/10.1038/s42254-023-00676-y.

[71] De Domenico M, Granell C, Porter MA, Arenas A. The physics of spreading processes in multilayer networks. Nature Phys 2016;12:901–6. https://doi.org/10.1038/nphys3865.

[72] De Domenico M. Multilayer Networks: Analysis and Visualization: Introduction to muxViz with R. Cham: Springer International Publishing; 2022. https://doi.org/10.1007/978-3-030-75718-2.

[73] Zitnik M, Leskovec J. Predicting multicellular function through multi-layer tissue networks. Bioinformatics 2017;33:i190–8. https://doi.org/10.1093/bioinformatics/btx252.

[74] Kumar T, Sethuraman R, Mitra S, Ravindran B, Narayanan M. MultiCens: Multilayer network centrality measures to uncover molecular mediators of tissue-tissue communication. PLoS Comput Biol 2023;19:e1011022. https://doi.org/10.1371/journal.pcbi.1011022.

[75] Wysocka M, Wysocki O, Zufferey M, Landers D, Freitas A. A systematic review of biologically-informed deep learning models for cancer: fundamental trends for encoding and interpreting oncology data. BMC Bioinformatics 2023;24:198. https://doi.org/10.1186/s12859-023-05262-8.

[76] Zhang X-M, Liang L, Liu L, Tang M-J. Graph Neural Networks and Their Current Applications in Bioinformatics. Front Genet 2021;12:690049. https://doi.org/10.3389/fgene.2021.690049.

[77] Guo D, Fan Y, Yue J-R, Lin T. A regulatory miRNA–mRNA network is associated with transplantation response in acute kidney injury. Hum Genomics 2021;15:69. https://doi.org/10.1186/s40246-021-00363-y.

[78] Lin Y, Wang L, Ge W, Hui Y, Zhou Z, Hu L, et al. Multi-omics network characterization reveals novel microRNA biomarkers and mechanisms for diagnosis and subtyping of kidney transplant rejection. J Transl Med 2021;19:346. https://doi.org/10.1186/s12967-021-03025-8.

[79] Perazzolo D, Fanton P, Barison I, Fedrigo M, Angelini A, Castellani C, et al. Improving Omics-Based Classification: The Role of Feature Selection and Synthetic Data Generation 2025. https://doi.org/10.48550/ARXIV.2505.03387.

[80] Morrissey CO. Unraveling the relationship between rejection and infection. The Journal of Heart and Lung Transplantation 2019;38:1233–4. https://doi.org/10.1016/j.healun.2019.06.008.

[81] Kanwar MK, Khush KK, Pinney S, Sherman C, Hall S, Teuteberg J, et al. Impact of cytomegalovirus infection on gene expression profile in heart transplant recipients. The Journal of Heart and Lung Transplantation 2021;40:101–7. https://doi.org/10.1016/j.healun.2020.11.008.

[82] Stewart S, Winters GL, Fishbein MC, Tazelaar HD, Kobashigawa J, Abrams J, et al. Revision of the 1990 Working Formulation for the Standardization of Nomenclature in the Diagnosis of Heart Rejection. The Journal of Heart and Lung Transplantation 2005;24:1710–20. https://doi.org/10.1016/j.healun.2005.03.019.

[83] Tibshirani R. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology 1996;58:267–88. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

[84] Fehlmann T, Kahraman M, Ludwig N, Backes C, Galata V, Keller V, et al. Evaluating the Use of Circulating MicroRNA Profiles for Lung Cancer Detection in Symptomatic Patients. JAMA Oncol 2020;6:714. https://doi.org/10.1001/jamaoncol.2020.0001.

[85] Barison I, Perazzolo D, Castellani C, Giarraputo A, Rossi E, Vedovelli L, et al. Transcriptomic profiling of cytomegalovirus infection in cardiac transplantation: proof-of-concept for a new strategy in tissue markers application. Front Immunol 2025;16:1581151. https://doi.org/10.3389/fimmu.2025.1581151.

[86] Bishop CM. Training with Noise is Equivalent to Tikhonov Regularization. Neural Computation 1995;7:108–16. https://doi.org/10.1162/neco.1995.7.1.108.

[87] Muthukrishnan R, Rohini R. LASSO: A feature selection technique in predictive modeling for machine learning. 2016 IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, India: IEEE; 2016, p. 18–20. https://doi.org/10.1109/ICACA.2016.7887916.

[88] Schölkopf B. Learning with kernels: support vector machines, regularization, optimization, and beyond. Cambridge, Mass: MIT Press; 2002.

[89] Lim M, Hastie T. Learning Interactions via Hierarchical Group-Lasso Regularization. Journal of Computational and Graphical Statistics 2015;24:627–54. https://doi.org/10.1080/10618600.2014.938812.

[90] Kramer O. Scikit-Learn. Machine Learning for Evolution Strategies, vol. 20, Cham: Springer International Publishing; 2016, p. 45–53. https://doi.org/10.1007/978-3-319-33383-0_5.

[91] Kaufman S, Rosset S, Perlich C, Stitelman O. Leakage in data mining: Formulation, detection, and avoidance. ACM Trans Knowl Discov Data 2012;6:1–21. https://doi.org/10.1145/2382577.2382579.

[92] Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, et al. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. Nucleic Acids Res 2015;43:W460–6. https://doi.org/10.1093/nar/gkv403.

[93] Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019;1:206–15. https://doi.org/10.1038/s42256-019-0048-x.

[94] Kim E, Kim J, Lee H, Kim S. Adaptive Data Augmentation to Achieve Noise Robustness and Overcome Data Deficiency for Deep Learning. Applied Sciences 2021;11:5586. https://doi.org/10.3390/app11125586.

[95] Lacan A, Sebag M, Hanczar B. GAN-based data augmentation for transcriptomics: survey and comparative assessment. Bioinformatics 2023;39:i111–20. https://doi.org/10.1093/bioinformatics/btad239.

[96] Li R, Wu J, Li G, Liu J, Xuan J, Zhu Q. Mdwgan-gp: data augmentation for gene expression data based on multiple discriminator WGAN-GP. BMC Bioinformatics 2023;24:427. https://doi.org/10.1186/s12859-023-05558-9.

[97] Spivey TL, Uccellini L, Ascierto ML, Zoppoli G, De Giorgi V, Delogu LG, et al. Gene expression profiling in acute allograft rejection: challenging the immunologic constant of rejection hypothesis. J Transl Med 2011;9:174. https://doi.org/10.1186/1479-5876-9-174.

[98] Bao Y, Wang L, Shi L, Yun F, Liu X, Chen Y, et al. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. Cell Mol Biol Lett 2019;24:38. https://doi.org/10.1186/s11658-019-0162-0.

[99] Wynn TA, Ramalingam TR. Mechanisms of fibrosis: therapeutic translation for fibrotic disease. Nat Med 2012;18:1028–40. https://doi.org/10.1038/nm.2807.

[100] Swaim AF, Field DJ, Fox-Talbot K, Baldwin WM, Morrell CN. Platelets Contribute to Allograft Rejection through Glutamate Receptor Signaling. The Journal of Immunology 2010;185:6999–7006. https://doi.org/10.4049/jimmunol.1000929.

[101] Wasowska BA, Qian Z, Cangello DL, Behrens E, Van Tran K, Layton J, et al. PASSIVE TRANSFER OF ALLOANTIBODIES RESTORES ACUTE CARDIAC REJECTION IN IgKO MICE1,2: Transplantation 2001;71:727–36. https://doi.org/10.1097/00007890-200103270-00007.

[102] Fenech A, Nicholls A, Smith FW. Indium (^{111} In)-labelled platelets in the diagnosis of renal transplant rejection: preliminary findings. BJR 1981;54:325–7. https://doi.org/10.1259/0007-1285-54-640-325.

[103] Di Francesco A, Fedrigo M, Santovito D, Natarelli L, Castellani C, De Pascale F, et al. MicroRNA signatures in cardiac biopsies and detection of allograft rejection. The Journal of Heart and Lung Transplantation 2018;37:1329–40. https://doi.org/10.1016/j.healun.2018.06.010.

[104] Peng Z, Duan Y, Zhong S, Chen J, Li J, He Z. RNA-seq analysis of extracellular vesicles from hyperphosphatemia-stimulated endothelial cells provides insight into the mechanism underlying vascular calcification. BMC Nephrol 2022;23:192. https://doi.org/10.1186/s12882-022-02823-6.

[105] Farzaneh-Far A. Vascular and valvar calcification: recent advances. Heart 2001;85:13–7. https://doi.org/10.1136/heart.85.1.13.

[106] Barabási A-L, Albert R. Emergence of Scaling in Random Networks. Science 1999;286:509–12. https://doi.org/10.1126/science.286.5439.509.

[107] Barabási A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–13. https://doi.org/10.1038/nrg1272.

[108] Parikshak NN, Gandal MJ, Geschwind DH. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. Nat Rev Genet 2015;16:441–58. https://doi.org/10.1038/nrg3934.

[109] O'Brien J, Hayder H, Zayed Y, Peng C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. Front Endocrinol 2018;9:402. https://doi.org/10.3389/fendo.2018.00402.

[110] Raichlin E, Edwards BS, Kremers WK, Clavell AL, Rodeheffer RJ, Frantz RP, et al. Acute Cellular Rejection and the Subsequent Development of Allograft Vasculopathy After Cardiac Transplantation. The Journal of Heart and Lung Transplantation 2009;28:320–7. https://doi.org/10.1016/j.healun.2009.01.006.

[111] Petrakopoulou P, Kübrich M, Pehlivanli S, Meiser B, Reichart B, Von Scheidt W, et al. Cytomegalovirus Infection in Heart Transplant Recipients Is Associated With Impaired Endothelial Function. Circulation 2004;110. https://doi.org/10.1161/01.CIR.0000138393.99310.1c.

[112] Wang J, Chen J, Sen S. MicroRNA as Biomarkers and Diagnostics. Journal Cellular Physiology 2016;231:25–30. https://doi.org/10.1002/jcp.25056.

[113] Duong Van Huyen J-P, Tible M, Gay A, Guillemain R, Aubert O, Varnous S, et al. MicroRNAs as non-invasive biomarkers of heart transplant rejection. European Heart Journal 2014;35:3194–202. https://doi.org/10.1093/eurheartj/ehu346.

[114] Long TA, Brady SM, Benfey PN. Systems Approaches to Identifying Gene Regulatory Networks in Plants. Annu Rev Cell Dev Biol 2008;24:81–103. https://doi.org/10.1146/annurev.cellbio.24.110707.175408.

[115] Berry GJ, Burke MM, Andersen C, Bruneval P, Fedrigo M, Fishbein MC, et al. The 2013 International Society for Heart and Lung Transplantation Working Formulation for the standardization of nomenclature in the pathologic diagnosis of antibody-mediated rejection in heart transplantation. The Journal of Heart and Lung Transplantation 2013;32:1147–62. https://doi.org/10.1016/j.healun.2013.08.011.

[116] Stewart S, Winters GL, Fishbein MC, Tazelaar HD, Kobashigawa J, Abrams J, et al. Revision of the 1990 Working Formulation for the Standardization of Nomenclature in the Diagnosis of Heart Rejection. The Journal of Heart and Lung Transplantation 2005;24:1710–20. https://doi.org/10.1016/j.healun.2005.03.019.

[117] Kivela M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. Journal of Complex Networks 2014;2:203–71. https://doi.org/10.1093/comnet/cnu016.

[118] De Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter MA, et al. Mathematical Formulation of Multilayer Networks. Phys Rev X 2013;3:041022. https://doi.org/10.1103/PhysRevX.3.041022.

[119] Yang Y, Fu X, Qu W, Xiao Y, Shen H-B. MiRGOFS: a GO-based functional similarity measurement for miRNAs, with applications to the prediction of miRNA subcellular localization and miRNA–disease association. Bioinformatics 2018;34:3547–56. https://doi.org/10.1093/bioinformatics/bty343.

[120] Zhao C, Wang Z. GOGO: An improved algorithm to measure the semantic similarity between gene ontology terms. Sci Rep 2018;8:15107. https://doi.org/10.1038/s41598-018-33219-y.

[121] Ruike Y, Ichimura A, Tsuchiya S, Shimizu K, Kunimoto R, Okuno Y, et al. Global correlation analysis for micro-RNA and mRNA expression profiles in human cell lines. J Hum Genet 2008;53:515–23. https://doi.org/10.1007/s10038-008-0279-x.

[122] Muniategui A, Nogales-Cadenas R, Vázquez M, L. Aranguren X, Agirre X, Luttun A, et al. Quantification of miRNA-mRNA Interactions. PLoS ONE 2012;7:e30766. https://doi.org/10.1371/journal.pone.0030766.

[123] Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Research 2020;48:D127–31. https://doi.org/10.1093/nar/gkz757.

[124] Vineetha S, Chandra Shekara Bhat C, Idicula SM. MicroRNA–mRNA interaction network using TSK-type recurrent neural fuzzy network. Gene 2013;515:385–90. https://doi.org/10.1016/j.gene.2012.12.063.

[125] Liang Z, Tan K, Yin Li C, Kuang Y. Self-feedback loop-containing synthetic mRNA switches for controlled microRNA sensing. Bioorganic Chemistry 2024;144:107081. https://doi.org/10.1016/j.bioorg.2023.107081.

[126] Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research 2000;28:27–30. https://doi.org/10.1093/nar/28.1.27.

[127] Philippidou D, Schmitt M, Moser D, Margue C, Nazarov PV, Muller A, et al. Signatures of MicroRNAs and Selected MicroRNA Target Genes in Human Melanoma. Cancer Research 2010;70:4163–73. https://doi.org/10.1158/0008-5472.CAN-09-4512.

[128] Siciliano V, Garzilli I, Fracassi C, Criscuolo S, Ventre S, Di Bernardo D. miRNAs confer phenotypic robustness to gene networks by suppressing biological noise. Nat Commun 2013;4:2364. https://doi.org/10.1038/ncomms3364.

[129] Ning H, Liu G, Li L, Liu Q, Huang H, Xie Z. Rational design of microRNA-responsive switch for programmable translational control in mammalian cells. Nat Commun 2023;14:7193. https://doi.org/10.1038/s41467-023-43065-w.

[130] Wang J, Liu S, Heallen T, Martin JF. The Hippo pathway in the heart: pivotal roles in development, disease, and regeneration. Nat Rev Cardiol 2018;15:672–84. https://doi.org/10.1038/s41569-018-0063-3.

[131] Vassalli G, Milano G, Moccetti T. Role of Mitogen-Activated Protein Kinases in Myocardial Ischemia-Reperfusion Injury during Heart Transplantation. Journal of Transplantation 2012;2012:1–16. https://doi.org/10.1155/2012/928954.

[132] Barac YD, Emrich F, Krutzwakd-Josefson E, Schrepfer S, Sampaio LC, Willerson JT, et al. The ubiquitin-proteasome system: A potential therapeutic target for heart failure. The Journal of Heart and Lung Transplantation 2017;36:708–14. https://doi.org/10.1016/j.healun.2017.02.012.

[133] Preiksaitis JK, Rosno S, Grumet C, Merigan TC. Infections Due to Herpesviruses in Cardiac Transplant Recipients: Role of the Donor Heart and Immunosuppressive Therapy. Journal of Infectious Diseases 1983;147:974–81. https://doi.org/10.1093/infdis/147.6.974.

[134] Dass D, Dhotre K, Chakraborty M, Nath A, Banerjee A, Bagchi P, et al. miRNAs in Herpesvirus Infection: Powerful Regulators in Small Packages. Viruses 2023;15:429. https://doi.org/10.3390/v15020429.

[135] Ribeiro AO, Schoof CRG, Izzotti A, Pereira LV, Vasques LR. MicroRNAs: modulators of cell identity, and their applications in tissue engineering. Microrna 2014;3:45–53. https://doi.org/10.2174/2211536603666140522003539.

[136] Gurtan AM, Sharp PA. The Role of miRNAs in Regulating Gene Expression Networks. Journal of Molecular Biology 2013;425:3582–600. https://doi.org/10.1016/j.jmb.2013.03.007.

[137] Ko DK, Brandizzi F. Network-based approaches for understanding gene regulation and function in plants. Plant J 2020;104:302–17. https://doi.org/10.1111/tpj.14940.

[138] Tonsho M, Michel S, Ahmed Z, Alessandrini A, Madsen JC. Heart transplantation: challenges facing the field. Cold Spring Harb Perspect Med 2014;4:a015636. https://doi.org/10.1101/cshperspect.a015636.

[139] Ding J, Ahangari F, Espinoza CR, Chhabra D, Nicola T, Yan X, et al. Integrating multiomics longitudinal data to reconstruct networks underlying lung development. Am J Physiol Lung Cell Mol Physiol 2019;317:L556–68. https://doi.org/10.1152/ajplung.00554.2018.

[140] Perazzolo D, Lazzaro G, Fiume A, Fanton P, Grisan E. Streamflow Forecasting: A Comparative Analysis of ARIMAX, Rolling Forecasting LSTM Neural Network and Physically Based Models in a Pristine Catchment. Water 2025;17:2341. https://doi.org/10.3390/w17152341.

[141] Perazzolo D, Castellani C, Grisan E. Uncovering Population PK Covariates from VAE-Generated Latent Spaces 2025. https://doi.org/10.48550/ARXIV.2505.02514.

[142] Lazzaro G, Basso S, Schirmer M, Botter G. Water management strategies for run-of-river power plants: Profitability and hydrologic impact between the intake and the outflow: Hydroeconomic Performances of Water Management Strategies. Water Resour Res 2013;49:8285–98. https://doi.org/10.1002/2013WR014210.

[143] Botter G, Basso S, Rodriguez-Iturbe I, Rinaldo A. Resilience of river flow regimes. Proc Natl Acad Sci USA 2013;110:12925–30. https://doi.org/10.1073/pnas.1311920110.

[144] Paparoditis E, Politis DN. The asymptotic size and power of the augmented Dickey–Fuller test for a unit root. Econometric Reviews 2018;37:955–73. https://doi.org/10.1080/00927872.2016.1178887.

[145] Hyndman RJ, Khandakar Y. Automatic Time Series Forecasting: The **forecast** Package for *R*. J Stat Soft 2008;27. https://doi.org/10.18637/jss.v027.i03.

[146] Ayzel G, Heistermann M. The effect of calibration data length on the performance of a conceptual hydrological model versus LSTM and GRU: A case study for six basins from the CAMELS dataset. Computers & Geosciences 2021;149:104708.

[147] Benvenuto D, Giovanetti M, Vassallo L, Angeletti S, Ciccozzi M. Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in Brief 2020;29:105340.

[148] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation 1997;9:1735–80. https://doi.org/10.1162/neco.1997.9.8.1735.

[149] Ali PJM, Faraj RH, Koya E, Ali PJM, Faraj RH. Data normalization and standardization: a technical report. Mach Learn Tech Rep 2014;1:1–6.

[150] Botchkarev A. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. arXiv Preprint arXiv:180903006 2018.

[151] Leavesley GH. Precipitation-runoff modeling system: User's manual. vol. 83. US Department of the Interior; 1984.

[152] U.S. Army Corps of Engineers. Hydrologic Modeling System HEC-HMS Technical Reference Manual. Hydrologic Engineering Center, Davis, CA; 2000.

[153] Hargreaves GH, Allen RG. History and evaluation of Hargreaves evapotranspiration equation. Journal of Irrigation and Drainage Engineering 2003;129:53–63.

[154] Hargreaves GH, Samani ZA. Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture 1985;1:96–9.

[155] Clark C. Storage and the unit hydrograph. Transactions of the American Society of Civil Engineers 1945;110:1419–46.

[156] Hallouin T. hydroeval: an evaluator for streamflow time series in Python 2021.

[157] McCuen RH, Knight Z, Cutter AG. Evaluation of the Nash–Sutcliffe efficiency index. Journal of Hydrologic Engineering 2006;11:597–602.

[158] Gupta HV, Kling H, Yilmaz KK, Martinez GF. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology 2009;377:80–91.

[159] Fabris L, Lazzaro G, Buddendorf WB, Botter G, Soulsby C. A general analytical approach for assessing the effects of hydroclimatic variability on fish habitat. Journal of Hydrology 2018;566:520–30.

[160] Knoben W, Freer J, Woods R. Technical note: inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. HydrolEarth Syst Sci 23: 4323–4331 2019.

[161] Sabzipour B, Arsenault R, Troin M, Martel J-L, Brissette F, Brunet F, et al. Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment. Journal of Hydrology 2023;627:130380.

[162] Dimri T, Ahmad S, Sharif M. Time series analysis of climate variables using seasonal ARIMA approach. Journal of Earth System Science 2020;129:1–16.

[163] Kaur J, Parmar KS, Singh S. Autoregressive models in environmental forecasting time series: a theoretical and application review. Environ Sci Pollut Res 2023;30:19617–41. https://doi.org/10.1007/s11356-023-25148-9.

[164] Rahajoe AD. Forecasting Feature Selection based on Single Exponential Smoothing using Wrapper Method. IJACSA 2019;10. https://doi.org/10.14569/IJACSA.2019.0100620.

[165] Liu Y, Wang H, Feng W, Huang H. Short Term Real-Time Rolling Forecast of Urban River Water Levels Based on LSTM: A Case Study in Fuzhou City, China. IJERPH 2021;18:9287. https://doi.org/10.3390/ijerph18179287.

[166] Giani G, Rico-Ramirez MA, Woods RA. A Practical, Objective, and Robust Technique to Directly Estimate Catchment Response Time. Water Resources Research 2021;57:e2020WR028201. https://doi.org/10.1029/2020WR028201.

[167] Yifru BA, Lim KJ, Lee S. Enhancing Streamflow Prediction Physically Consistently Using Process-Based Modeling and Domain Knowledge: A Review. Sustainability 2024;16:1376. https://doi.org/10.3390/su16041376.

[168] Troin M, Arsenault R, Wood AW, Brissette F, Martel J-L. Generating ensemble streamflow forecasts: A review of methods and approaches over the past 40 years 2021.

[169] Waqas M, Humphries UW. A critical review of RNN and LSTM variants in hydrological time series predictions. MethodsX 2024;13:102946. https://doi.org/10.1016/j.mex.2024.102946.

[170] Lin Y, Wang D, Wang G, Qiu J, Long K, Du Y, et al. A hybrid deep learning algorithm and its application to streamflow prediction. Journal of Hydrology 2021;601:126636. https://doi.org/10.1016/j.jhydrol.2021.126636.

[171] Schnider TW, Minto CF, Cambus PL, Andresen C, Goodale DB, Shafer SL, et al. The Influence of Method of Administration and Covariates on the Pharmacokinetics of Propofol in Adult Volunteers. Anesthesiology 1998;88:1170–82.

[172] Sherwin CMT, Kiang TKL, Spigarelli MG, Ensom MHH. Fundamentals of Population Pharmacokinetic Modelling: Validation Methods. Clinical Pharmacokinetics 2012;51:573–90.

[173] Hutmacher MM, Kowalski KG. Covariate Selection in Pharmacometric Analyses: A Review of Methods. British Journal of Clinical Pharmacology 2015;79:132–47.

[174] Kingma DP, Welling M, others. An introduction to variational autoencoders. Foundations and Trends® in Machine Learning 2019;12:307–92.

[175] Gomari DP, Schweickart A, Cerchietti L, Paietta E, Fernandez H, Al-Amin H, et al. Variational Autoencoders Learn Transferrable Representations of Metabolomics Data. Communications Biology 2022;5:645.

[176] Ribbing J, Nyberg J, Caster O, Jonsson EN. The LASSO—a Novel Method for Predictive Covariate Model Building in Nonlinear Mixed Effects Models. Journal of Pharmacokinetics and Pharmacodynamics 2007;34:485–517.

[177] Gamal A, Sharafeldeen A, Alnaghy E, Alghandour R, Alghamdi NS, Ali KM, et al. A Novel Machine Learning Approach for Predicting Neoadjuvant Chemotherapy Response in Breast Cancer: Integration of Multimodal Radiomics with Clinical and Molecular Subtype Markers. IEEE Access 2024.

[178] Venkataramanan R, Swaminathan A, Prasad T, Jain A, Zuckerman S, Warty V, et al. Clinical Pharmacokinetics of Tacrolimus. Clinical Pharmacokinetics 1995;29:404–30.

[179] Chen D, Yao Q, Chen W, Yin J, Hou S, Tian X, et al. Population PK/PD Model of Tacrolimus for Exploring the Relationship Between Accumulated Exposure and Quantitative Scores in Myasthenia Gravis Patients. CPT: Pharmacometrics & Systems Pharmacology 2023;12:963–76. https://doi.org/10.1002/psp4.12966.

[180] Nanga TM, Doan TTP, Marquet P, Musuamba FT. Toward a Robust Tool for Pharmacokinetic-Based Personalization of Treatment with Tacrolimus in Solid Organ Transplantation: A Model-Based Meta-Analysis Approach. British Journal of Clinical Pharmacology 2019;85:2793–823.

[181] Kingma DP, Welling M. An Introduction to Variational Autoencoders. FNT in Machine Learning 2019;12:307–92. https://doi.org/10.1561/2200000056.

[182] Muthukrishnan R, Rohini R. LASSO: A feature selection technique in predictive modeling for machine learning. 2016 IEEE International Conference on Advances in Computer

Applications (ICACA), Coimbatore, India: IEEE; 2016, p. 18–20. https://doi.org/10.1109/ICACA.2016.7887916.

[183] Kramer O. Scikit-Learn. Machine Learning for Evolution Strategies, vol. 20, Cham: Springer International Publishing; 2016, p. 45–53. https://doi.org/10.1007/978-3-319-33383-0_5.

[184] Jones H, Rowland-Yeo K. Basic Concepts in Physiologically Based Pharmacokinetic Modeling in Drug Discovery and Development. CPT Pharmacom & Syst Pharma 2013;2:1–12. https://doi.org/10.1038/psp.2013.41.

[185] Svensson RJ, Jonsson EN. Efficient and relevant stepwise covariate model building for pharmacometrics. CPT Pharmacom & Syst Pharma 2022;11:1210–22. https://doi.org/10.1002/psp4.12838.

[186] Staatz CE, Tett SE. Clinical Pharmacokinetics and Pharmacodynamics of Tacrolimus in Solid Organ Transplantation: Clinical Pharmacokinetics 2004;43:623–53. https://doi.org/10.2165/00003088-200443100-00001.

[187] Størset E, Holford N, Midtvedt K, Bremer S, Bergan S, Åsberg A. Importance of hematocrit for a tacrolimus target concentration strategy. Eur J Clin Pharmacol 2014;70:65–77. https://doi.org/10.1007/s00228-013-1584-7.

[188] Deininger KM, Vu A, Page RL, Ambardekar AV, Lindenfeld J, Aquilante CL. *CYP 3A* pharmacogenetics and tacrolimus disposition in adult heart transplant recipients. Clinical Transplantation 2016;30:1074–81. https://doi.org/10.1111/ctr.12790.

[189] Paschier A, Destere A, Monchaud C, Labriffe M, Marquet P, Woillard J. Tacrolimus population pharmacokinetics in adult heart transplant patients. Brit J Clinical Pharma 2023;89:3584–95. https://doi.org/10.1111/bcp.15857.

Appendix

During my PhD, I actively contributed to two additional research projects. The first study focused on the transcriptomic profiling of Cytomegalovirus (CMV) infection in endomyocardial biopsies from heart transplant recipients, proposing a novel molecular strategy to distinguish infection from acute cellular rejection. The second was an international collaborative project with the International Society for Heart and Lung Transplantation (ISHLT), aimed at analyzing global practices in the collection, processing, and biobanking of explanted hearts. My involvement in both studies included study design, statistical analysis, data interpretation, and manuscript preparation.

Transcriptomic Signature of CMV Infection in Heart Transplantation

Endomyocardial biopsy (EMB) remains the gold standard for monitoring graft status in heart transplant recipients, yet distinguishing between cytomegalovirus (CMV) infection and acute cellular rejection (ACR) can be particularly challenging due to their similar histopathological features. To address this diagnostic limitation, we conducted a translational research study aimed at identifying a transcriptomic signature capable of differentiating CMV-related inflammation from immune-mediated rejection in EMB specimens. This proof-of-concept study aimed to identify a distinct transcriptomic signature associated with CMV-related inflammation, offering potential tissue biomarkers to support EMB interpretation. The study included mRNA and microRNA (miRNA) expression profiling on formalin-fixed, paraffin-embedded (FFPE) EMB specimens from three patient groups: CMV infection (n=6), ACR (n=5), and histologically normal controls (n=5). High-density microarrays were used to quantify transcript levels, followed by extensive bioinformatic analysis to identify differentially expressed genes (DEGs) and miRNAs. In the infection vs rejection comparison, we identified 18 DEGs and 12 significantly deregulated miRNAs. Among these, two mRNAs, IL7R and GZMK, and two miRNAs: miR-93-5p and miR-345-5p, showed the most discriminative expression patterns. Both genes and miRNAs were markedly downregulated in CMV-infected samples compared to ACR, suggesting their potential role in distinguishing between these two inflammatory conditions. To explore the diagnostic potential of these markers, we developed a machine learning classifier using a kernel-based support vector machine (KSVM) and achieved a cross-validated classification accuracy of 91% with an area under the ROC curve (AUC) of 0.87. Further integrative analysis explored the regulatory interaction between miRNAs and mRNAs, revealing that miR-93-5p targets several genes involved in the IL-7 signaling and GZMK-related pathways, supporting a mechanistic link between molecular regulation and immune activation in infection. This study demonstrates the feasibility of using transcriptomic profiling to support EMB-based diagnosis and paves the way for future applications of molecular tissue markers in transplant pathology. My personal contribution included the interpretation of transcriptomic data, design of bioinformatic and statistical analyses, development of the classification pipeline, and manuscript writing. The results of this research were published in Frontiers in Immunology in May 2025 under the title “Transcriptomic profiling of cytomegalovirus infection in cardiac transplantation: proof-of-concept for a new strategy in tissue markers application.”

Biobanking Practices for Explanted Hearts: An ISHLT Survey

To better understand and harmonize global practices in the handling of explanted hearts, we participated in an international collaborative study promoted by the International Society for Heart and Lung Transplantation (ISHLT), aimed at assessing current protocols for gross examination, tissue sampling, and storage of explanted cardiac grafts. The study was based on a structured survey distributed to transplant centers across Europe, North America, and other regions, and collected responses regarding grossing protocols, the number and type of tissue samples collected (both fresh and formalin-fixed), coronary artery sampling, storage duration, and the use of standardized reporting forms. The analysis revealed significant heterogeneity among centers. While the majority of centers (78.8%) reported adopting a grossing protocol, North America had the highest adoption rates and the most extensive sampling procedures, often collecting 10–25 tissue samples per heart. In contrast, centers in Europe and other countries more frequently collected fewer samples, particularly fresh tissue. Coronary artery sampling and long-term storage practices also varied, with only a portion of centers preserving hearts indefinitely. Importantly, fresh tissue collection, crucial for subsequent analysis, was inconsistently applied and often limited by institutional infrastructure. The results underscore the need for standardized international protocols to improve the diagnostic and scientific value of explanted hearts, enabling more consistent clinical evaluation and supporting advanced research. The study advocates for greater involvement of pathologists in biobanking workflows and calls for unified guidelines on sample type, volume, and storage methods to enhance data comparability and foster innovation in transplant pathology. My contribution to this study focused on performing the statistical analysis of survey data, generating visual representations of results, and revising the manuscript. The work was published in JHLT Open in 2025 under the title “The value of systematic study and biobanking of explanted hearts: Insights from an international ISHLT pathology survey.”

frontiers | Frontiers in Immunology

TYPE Perspective
PUBLISHED 16 May 2025
DOI 10.3389/fimmu.2025.1581151

OPEN ACCESS

EDITED BY
Martin Kauke-Navarro,
Yale-New Haven Hospital, United States

REVIEWED BY
Javier Carbone,
Gregorio Marañón Hospital, Spain
Zhendong Ding,
Central South University, China

*CORRESPONDENCE
Annalisa Angelini
annalisa.angelini@unipd.it

RECEIVED 21 February 2025
ACCEPTED 09 April 2025
PUBLISHED 16 May 2025

CITATION
Barison I, Perazzolo D, Castellani C, Giarraputo A, Rossi E, Vedovelli L, Minuzzo SA, Tessari C, Pradegan N, Toscano G, Tona F, Basso C, Gerosa G, Mandruzzato S, Abate D, Gregori D, Angelini A and Fedrigo M (2025) Transcriptomic profiling of cytomegalovirus infection in cardiac transplantation: proof-of-concept for a new strategy in tissue markers application. *Front. Immunol.* 16:1581151. doi: 10.3389/fimmu.2025.1581151

Transcriptomic profiling of cytomegalovirus infection in cardiac transplantation: proof-of-concept for a new strategy in tissue markers application

Ilaria Barison[1], Diego Perazzolo[1], Chiara Castellani[1], Alessia Giarraputo[1], Elisabetta Rossi[2,3], Luca Vedovelli[4], Sonia Anna Minuzzo[2], Chiara Tessari[5], Nicola Pradegan[5], Giuseppe Toscano[5], Francesco Tona[6], Cristina Basso[1], Gino Gerosa[4], Susanna Mandruzzato[2,3], Davide Abate[7], Dario Gregori[4], Annalisa Angelini[1*] and Marny Fedrigo[1]

[1]Cardiovascular Pathology, Department of Cardiac, Thoracic and Vascular Sciences and Public Health, University of Padova, Padova, Italy, [2]Department of Surgery, Oncology and Gastroenterology, Oncology Section, University of Padova, Padova, Italy, [3]Immunology and Molecular Oncology Diagnostics, Veneto Institute of Oncology (IOV) - Istituto di Ricovero e Cura a Carattere Scientifico, Padova, Italy, [4]Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, Vascular Sciences, and Public Health, University of Padova, Padova, Italy, [5]Division of Cardiac Surgery, Department of Cardiac, Thoracic and Vascular Sciences and Public Health, University of Padova, Padova, Italy, [6]Cardiology Unit, Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Padova, Italy, [7]Department of Molecular Medicine, University of Padova, Padova, Italy

Background: Cytomegalovirus (CMV) infection is a relevant threat to heart-transplanted patients during the first year after surgery, leading to increased morbidity and, in some cases, mortality. This *proof-of-concept* study aims to assess the transcriptomic profile of CMV infection in cardiac transplanted patients as a new diagnostic approach to discriminate infection and Acute Cellular Rejection (ACR) on EMB specimens.

Methods: We performed a microarray-based messenger RNA (mRNA) and micro-RNA (miRNA) profiling. We analyzed three patient groups in the setting of CMV viremia and inflammatory infiltrate: a control group (n=5), an ACR group (n=5), and an infection group (n=6). Differentially expressed mRNA and miRNA were further investigated through bioinformatic pathway analysis.

Results: Focusing on infection vs rejection comparison, we investigated the role of the 18 differentially expressed mRNAs and the 12 miRNAs with the most significative p-value (gene level fold change, FC <-2 or >2, p-value <0.05). Based on the bioinformatic analysis, we explored the regulatory effects of these miRNAs on the mRNA pathways independently identified in the same samples. The results showed that two genes, IL7R and GZMK (-38.63 and -3.15 FC, respectively), and two miRNAs, mir-93-5p and mir-345-5p (-2.63 and -2.18 FC, respectively), are differentially expressed in infection and can be exploited to differentiate CMV-positive from ACR-positive EMB specimens, reaching an AUC of 0.87 and an accuracy of 91% at cross-validation.

10.3389/fimmu.2025.1581151

Conclusions: We have identified a distinctive combined molecular profile of mRNAs and miRNAs for infection in post-cardiac transplant follow-up. Based on IL7R, GZMK, mir-93-5p, and mir-345-5p we suggest a novel possible workflow to distinguish infection, where those markers are downregulated, from rejection, where they are overexpressed, on EMB specimens. This analysis showed good accuracy and promising predictive performance. The future combined analysis of these genes and these miRNAs through user-friendly techniques, such as quantitative PCR, could reduce turn-around time and improve our diagnostic power for distinguishing CMV infection from ACR in EMB specimens.

KEYWORDS

cytomegalovirus, heart transplant, rejection, transcriptomic profiling mRNA, miRNA, biomarkers

1 Introduction

Cytomegalovirus (CMV) is the most clinically relevant post-transplant infectious agent. CMV is a member of the beta-Herpesviridae family that, in the normal population, latently infects 50-90% of individuals, but normally takes an asymptomatic course (1). Differently, in cardiac transplanted patients, CMV infection impacts morbidity and mortality (2, 3), and some studies showed that CMV infection is a risk factor for Antibody Mediated Rejection (AMR) and Cardiac Allograft Vasculopathy (CAV) (4–7).

The cardiac transplanted patients can be stratified according to the donor/recipient (D/R) serological status. The highest risk is associated with the mismatch between donor-positive and recipient-negative (D+/R-). Based on the approach chosen by each transplant center, these patients may undergo pre-emptive therapy (based on the antiviral administration for early asymptomatic CMV viremia detected by surveillance testing) or antiviral prophylaxis (2, 8, 9).

Routine EMB tissue exhamination represents a crucial procedure to determine CMV replication on the graft and thus intervene with a proper antiviral therapy and taper the immunosuppression regimen accordingly. However, the discrimination between Acute Cellular Rejection (ACR) and the inflammation process triggered by CMV infection itself is quite cryptic. In case of suspected CMV infection, both nucleic acid and immunohistochemical tests should be performed on EMB specimens, e.g. Immunohistochemical staining on Formalin Fixed Paraffine Embedded (FFPE) EMB slices and Polymerase Chain Reaction with the nucleic acid extracted from EMB tissue. Those tests are designed to detect viral components, such as the CMV genome or proteins, in tissue samples. Still, these tests are prone to false results due to CMV activation's patchy nature in tissue, and the EMB procedure sampling errors.

Seeking novel biomarkers associated with rejection monitoring has been investigated over years in tissue and liquid biopsy (10–13). Transcriptomic studies opened new horizons on cardiac rejection allograft monitoring, showing that Rejection Associated Transcript analyses are similar through solid organ transplants and can differentiate among Cellular Rejection, Antibody-Mediated Rejection, and no-rejecting EMBs (11, 14, 15). Unfortunately, most of these studies focused mainly on inflammation but not on the infection. This limitation urges the need for new approaches that could help the pathologist in characterizing the allograft's status, differentiating the inflammation itself from the virus induced inflammation pattern.

In this *proof-of-concept* study, we investigate the transcriptomic profile of CMV-positive patients at mRNA and miRNA levels to determine new candidate biomarkers to distinguish inflammatory infiltrate caused by infection from that due to rejection on FFPE EMB specimens. For this purpose, we analyzed the mRNA and miRNA profile of FFPE EMB through high-density Clariom S Affymetrix GeneChip arrays (Thermofisher Scientific, USA) platform. We demonstrated that combining different biomarkers can significantly improve CMV infection detection and discriminate it from ACR during EMB pathological assessment.

2 Materials and methods

All the details about the Study design, patient selection, EMBs histological evaluation, RNA extraction protocol, Microarray-based mRNA and miRNA analysis protocols, bioinformatics, and statistical analysis are provided in the Extended Methods in Supplementary Materials.

3 Results

3.1 Study population characteristics

The patients enrolled in this study were divided into three groups according to the ACR grade and the CMV viremia. Recipients' and Donors' characteristics are shown in Table 1. Among selected patients, 68.75% were male, and the mean age

10.3389/fimmu.2025.1581151

TABLE 1 Study population characteristics.

Features	Control (n=5)	Rejection (n=5)	Infection (n=6)	t-test p-value Control vs Rejection	t-test p-value Control vs Infection	t-test p-value Rejection vs Infection
Recipient Age (y)	61 ± 5.33	61.6 ± 9.97	54.3 ± 19.69	0.91	0.45	0.46
Recipient gender (male, %)	60	60	100	1	0.17	0.17
Ischemic Heart Disease (%)	60	40	66.7	0.57	0.84	0.43
Dilated Cardiomyopathy (%)	20	40	66.7	0.57	0.84	0.43
Arrhythmogenic Right Ventricular Cardiomyopathy (%)	20	0	16.65	0.37	0.90	0.36
Obstructive Cardiomyopathy (%)	0	20	0	0.37	–	0.37
Donor Age (y)	51 ± 20.29	64 ± 6.67	43 ± 17.67	0.76	0.52	0.22
Donor gender (male, %)	80	60	33.3	0.75	0.87	0.88
Cold Ischemia Time (m)	184.2 ± 49.88	195.4 ± 52.04	189.83 ± 61.33	0.75	0.87	0.88
Circulating DSA at the time of EMB (d)	0	0	33.3	–	0.18	0.18
Time between HTx and EMB (d)	207 ± 113.90	62 ± 10.86	77 ± 79.00	0.046	0.067	0.668
Neutrophils (10^9/l)	4.10 ± 2.12	4.16 ± 1.77	5.24 ± 3.37	0.96	0.51	0.52
Lymphocytes (10^9/l)	0.76 ± 0.40	0.69 ± 0.21	0.99 ± 0.78	0.75	0.56	0.41
P-AST (U/l)	46.4 ± 52.92	23.6 ± 4.92	22.4 ± 15.48	0.39	0.37	0.87
P-ALT (U/l)	25.8 ± 20.59	21.8 ± 10.14	24.4 ± 27.07	0.71	0.51	0.37
P-γGT (U/l)	24.2 ± 9.67	58.6 ± 47.86	94 ± 67.50	0.22	0.08	0.38
e-GFR (µmol/l)	42 ± 15.76	75 ± 11.62	63.2 ± 41.23	0.007	0.33	0.56
Creatinine (µmol/l)	146.6 ± 30.37	86.4 ± 13.30	149.5 ± 133.77	0.0077	0.96	0.30
***Cyclosporine* (µg/l)**	135 ± 57.05	172.2 ± 36.25	167.83 ± 48.43	0.27	0.33	0.87
Ejection Fraction (%)	63	59	58.83	0.45	0.97	0.22

Age, sex, and the main biochemical parameters are reported as mean ± standard deviation. Y, Years; d, days; m, minutes; HTx, Heart Transplantation; eGFR, estimated Glomerular Filtration Rate. The bold style was used to highlight the features reported in the first column from the numerical values reported in the other cells of the table.

was 61. Ischemic Heart Disease was the most common diagnosis leading to transplant. These patients were homogenous for Circulating DSA at the time of EMB, lymphocyte count, hepatic functionality, and Ejection Fraction. The control group showed slightly worse renal functionality and a higher time interval between heart transplant and EMB.

3.2 Microarray transcriptomic analysis of FFPE EMBs

We analyzed the transcriptomic profile of our samples through a microarray-based platform. The bioinformatic analysis compared the expression levels of detected mRNAs in the Controls, the Rejection group, and the Infection group.

Notably, 293 genes were differentially expressed between Control and Infection groups (82 up-regulated and 211 down-regulated), 407 genes between Control and Rejection groups (126 up-regulated and 281 down-regulated), and 18 genes between Infection and Rejection groups (10 up-regulated and 8 down-regulated) (p-value <0.05).

Since we aimed to define a possible molecular signature to distinguish infection from rejection at the tissue level, we focused our attention on the differentially expressed genes (DEG) that emerged in the comparison between these two study groups. In particular, PHLDB2, HP1BP3, HMGCS2, UBC, LDB3, ALPK2,

MYH7B, ACADVL, HSPB6, and TBX20 resulted to be overexpressed in Infection compared to Rejection, while PDCL2, OCRL, TAOK3, PLAC8, GZMK, MNDA, IL7R, P2RY14 were downregulated in the same comparison ($p<0.05$) (Figure 1A). Afterward, we investigated the role of all these DEGs. We consulted several databases, such as the GeneCards database and WikiPathways TAC tool, focusing on their cellular functions and the biological processes they are involved in. This analysis revealed that these DEGs are mainly involved in cardiomyocyte metabolism (e.g. HMGCS2, ACADVL), IL-7 signaling and regulatory circuits of the STAT3 signaling (e.g. IL7R), DNA replication (e.g. HP1BP3) and cardiac progenitor differentiation pathways (e.g. PHLDB2, LDB3, ALPK2, TBX20) (Supplementary Table S1).

As shown in the Volcano Plot (Figure 1A), all the significantly differently expressed genes are quite homogeneous in terms of fold-change values, except for GZMK and IL7R. In detail, IL7R resulted downregulated in the Infection group compared to the Rejection (-38.63 fold-change; 12.69 (log2) Avg in the Rejection group vs 7.41 (log2) Avg in the Infection group). IL7R is highly expressed in conventional mature T-cells, except for regulatory T-cells that show a low level of IL7R on their surface (16, 17). Once activated, IL7R induces proliferative and anti-apoptotic signals mainly by activating several cellular pathways, such as JAK/STAT, PI3K/Akt, and MAPK/ERK pathways (18, 19). In addition, IL7R has been reported to activate chronic inflammation (20), autoimmunity (21), and allograft rejection in rodent models (22, 23). Likewise, GZMK is downregulated in the infection group (-3.15 fold-change; 9.27 (log2) Avg in Rejection group vs 7.62 (log2) Avg in Infection group). GZMK is a member of serine proteases that preferentially cleave after basic residues (24, 25), and it is expressed in different immune cell types, such as NK cells, cytotoxic T cells (26), and macrophages (27). GZMK is also reported to be involved in several pathologic conditions, such as cardiovascular diseases (28), vasculitis (29), and endothelial activation, promoting the release of proinflammatory cytokines from endothelial cells (30).

The heatmap reported in Figure 1B confirms that the DEG values are homogenous within each group, in terms of the expression level. Moreover, the hierarchical clustering correctly classifies nine out of eleven samples (82%), in agreement with histology and CMV serology. Only two samples were assigned to the wrong group: sample #36, selected as Infection but classified as Rejection by hierarchical clustering, and sample #38, selected as Rejection but classified as Infection by hierarchical clustering. In detail, the misclassification of these samples is linked to the values of GZMK and IL7R genes. In our cohort, the expression levels of these two genes were generally high in the Rejection group and low in the Infection group. Sample #36 showed a low level of expression of IL7R and GZMK genes, and the system classified it as "Infection" despite it being negative for the CMV serology test. On the other hand, sample #38 showed a high level of expression for IL7R and GZMK genes, and the system classified it as "Rejection" even though the CMV serology positivity. This misclassification suggests that those genes are crucial in differentiating Infections, with low expression, from Rejections, with high expression.

3.3 Microarray miRNAs analysis of FFPE EMB

The same total RNA extracted and analyzed for mRNA profiling was evaluated for miRNA profiling. The Clariom-S miRNA 4.0 chip tested the expression level of 2578 different miRNAs for each patient, and we performed the same comparisons conducted for the mRNA analysis (gene level fold-change <-2 or >2, p-value<0.05). The statistical evaluation of the raw data revealed that 513 miRNAs were differentially expressed between the Control and Rejection groups, with 477 miRNAs upregulated and 36 downregulated, while 111 (62 upregulated and 49 downregulated) between the Control and Infection groups. The comparison between the Infection and the Rejection groups showed that 386 miRNAs were differentially expressed (32 upregulated and 354 downregulated). Among these 386 miRNAs, 29 emerged differentially expressed only in the latter comparison. We further filtered out those results, focusing on the miRNA with the highest p-value. Hence we selected twelve hsa-miRNAs that showed the highest p-value in infection versus vs rejection comparison: miR-8075 (fold-change= 2.23; p-value= 0.0052), miR-93-5p (fold-change= -2.63; p-value= 0.0058), miR-3651 (fold-change= -2.26; p-value= 0.0059), miR-345-5p (fold-change= -2.25; p-value= 0.0061), miR-6722 (fold-change= 2.57; p-value= 0.0098), miR-1296-5p (fold-change= -2.08; p-value= 0.0112), miR-3162-5p (fold-change= -2.15; p-value= 0.0114); miR-532-5p (fold-change= -3.26; p-value= 0.0177), miR-4433-3p (fold-change= -2.02; p-value= 0.0293), miR-6782-5p (fold-change= -2.08; p-value= 0.0365), miR-342-5p (fold-change= -2.18; p-value=0.0371), and miR-210-3p (fold-change= -2.54; p-value= 0.0469) (Figure 1C).

Once determined our miRNAs of interest, we explored their function through a literature review. The screening of PubTator, GeneCards, and PubMed databases revealed that these miRNAs are mainly involved in PI3K/Akt, STAT, and HIF1α pathways, regulating cell proliferation and apoptosis, along with mitochondria stress response.

3.4 Analysis of mRNA and miRNA interactions in CMV infection

Once we identified the mRNAs and miRNAs differentially expressed in the Infection vs Rejection groups, we further investigated in silico the possible interactions among mRNAs and the target genes of miRNAs.

We consulted TAC and Microcosm databases to define the target genes of the selected miRNAs. Through a symmetric matrix (Supplementary Figure S2), we highlighted the number of shared genes among each couple of entries. In detail, it showed that only miR-93-5p targets some of the mRNAs identified as DEG in the Infection group.

Since we focused on the miRNAs shared target genes, we saw that the most relevant miRNAs were miR-93-5p, miR-345-5p, miR-532-5p, miR-342-5p, and miR-210-3p. Indeed, these miRNAs shared a higher number of target genes: miR-93-5p shared 101 target genes with miR-345-5p, 100 target genes with miR-532-5p,

10.3389/fimmu.2025.1581151

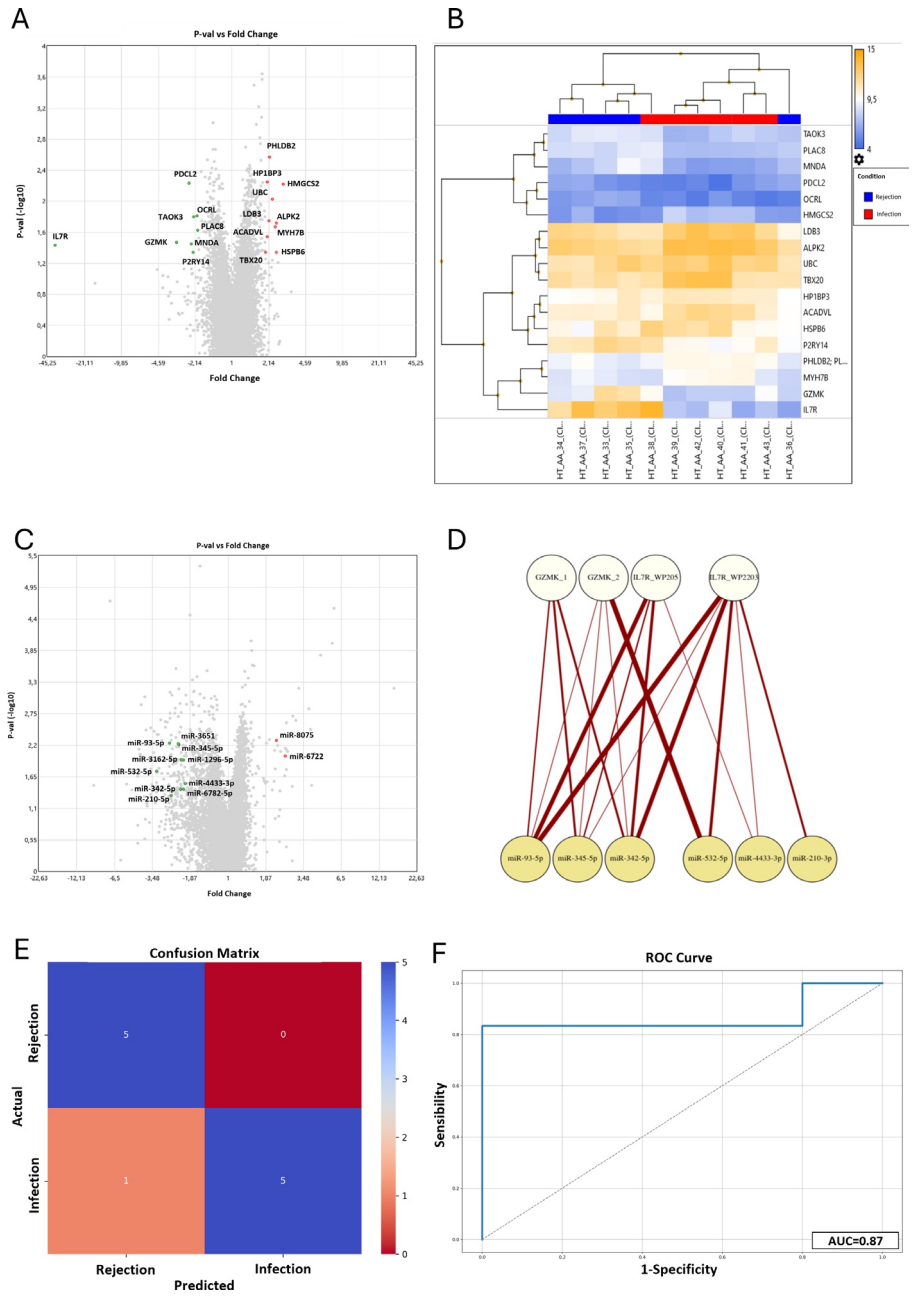


FIGURE 1
Results of the microarray mRNA analysis in the Infection vs Rejection group. **(A)** The Volcano Plot reported the differentially expressed genes between the Infection and Rejection groups. On the horizontal axis is reported the fold-change value (fold-change <-2 and >2), while the vertical axis reported the p-values (<0.05) of the same genes. **(B)** The Hierarchical Clustering reported the fluorescence level of each DEG (higher level in orange and lower level in blue). In the upper part of the graph is reported the clustering of the patients in the two study groups, Rejection in blue and Infection in red. **(C)** Results of the microarray miRNAs analysis in the Infection vs Rejection group. The Volcano Plot reported the differentially expressed miRNAs between the Infection and Rejection groups. On the horizontal axis is reported the fold-change value (fold-change <-2 and >2) while the vertical axis reports the p-values (<0.05) of the same miRNAs. **(D)** The bipartite graph between GZMK and IL7R pathways and miRNAs target genes, obtained by igraph library in R. The upper white nodes layer represents the Cytoscape pathways selected: GZMK glucose metabolic (GZMK_1), GZMK type I immune response (GZMK_2), IL7R WP205 - IL-7 signaling pathway (IL7R_WP205), and L7R WP2203 - Thymic stromal lymphopoietin signaling pathway (IL7R_WP2203). The lower yellow node layer represents the infection-related miRNAs. **(E)** Confusion Matrix. The panel shows the performance of the Support Vector Machine model (SVM) to classify infection and rejection at cross-validation. The y-axis represents the actual enrolled group (5 samples for rejection, and 6 samples for the infection group). The x-axis represents the samples classified by the SVM. **(F)** ROC curve analysis: Area Under the Curve (AUC) at cross-validation for infection and rejection discrimination is reported.

78 target genes with miR-342-5p, and 66 target genes with miR-210-3p (Supplementary Figure S2A). Next, we represented these connections with an undirected weighted graph: the thickness of the edges is proportional to the number of targets shared between two miRNAs (Supplementary Figure S2B). Once again, the miRNAs that were more relevant were miR-93-5p, miR-345-5p, miR-532-5p, miR-342-5p, and miR-210-3p.

Interestingly, miR-93-5p targets some mRNAs identified in infection vs rejection: ACADVL, ALPK2, OCRL, PDCL2, and UBC. Using the EnrichR tool, we saw that these genes are mainly involved in fatty acid metabolism pathways (Supplementary Figures S2C, D).

3.5 The interplay among miRNA targets and GZMK and IL7R pathways

The analysis of mRNAs revealed that GZMK and IL7R are crucial for distinguishing infection from rejection. Thus, we investigated the interplay between the gene pathways and our selected miRNAs. Consulting the Cytoscape database, we chose the two most significant pathways for GZMK and IL7R based on the p-value associated. The genes reported to be involved in these pathways were matched with the miRNAs' target genes to identify the shared ones, and the results are shown in (See Supplementary Materials, Supplementary Table S2). Notably, only seven miRNAs were shown to target some genes involved in GZMK and IL7R pathways, and their interplay is graphically represented by the bipartite graph reported in Figure 1D. Once again, miR-93-5p shared more target genes with the pathways analyzed, especially with IL7R pathways, such as MAPK1, JAK1, MYC, PIK3R2, STAT3, and CRLF2. MiR-345-5p and miR-342-5p target several genes in GZMK and IL7R pathways, such as ADHFE1, PTK2B, IRF1, PIK3R2, CRLF2, MAP2K2, MAPK3, LCK.

Thus, miR-93-5p and miR-345-5p showed the strongest connections with IL7R and GZMK pathways. Both of them were up-regulated in the Rejection group: mir-93-5p showed 5.98 Avg (log2) in the Rejection group vs 4.59 Avg (log2) in the Infection group; mir-342-5p showed 2.18 Avg (log2) in the Rejection group vs 1.6 Avg (log2) in the Infection group. Li et al. demonstrated that the miR-93 family takes part in PTEN regulation (31). PTEN is a tumor suppressor mutated in many cancers that antagonizes the PI3K/Akt pathway, modulating cell cycle progression and cell survival (32, 33). Furthermore, several studies demonstrated that this miRNA plays an important role in enhancing endothelial activities (34) and regulating integrin-β8 expression (35). Additionally, miR-345-5p showed several shared target genes with other miRNAs. This miRNA is reported to play a role in the regulation of HIF1α, and consequently, modulation of TGFβ/Smad2/Smad3 signaling (36), linked to ischemic damage response. Liu and colleagues reported that miR-345-5p can downregulate the TLR4/NF-κB pathway, altering inflammatory response and apoptosis (37).

Taken together, all these results showed that miR-93-5p and miR-345-5p are the best-ranked candidates to distinguish infection from rejection since they are down-regulated in CMV-positive patients and up-regulated in the Rejection group.

3.6 Performance assessment of the combined molecular markers

We are aware that our study population is small, and this leads to a reduced statistical power. This was confirmed by the *Post Hoc* Power analysis, which revealed that our cohort is associated with a value of (1-β) of 0.19, and, consequently, with a high risk of false negative results.

However, once defined our new panel of combined markers, we assessed the accuracy of this novel approach to distinguish infection from rejection. We implemented a classifier based on the identified markers using a Support Vector Machine with a polynomial kernel (KSVM). We then performed stratified cross-validation to evaluate its performance. As shown in the Confusion Matrix (Figure 1E), based on the level of expression of GZMK, IL7R, mir-93-5p, and mir-345-5p, the model correctly classified as "rejection" all the real rejection samples. On the other hand, the model correctly classified five out of six infection samples (83,3%), with only one misclassification. This performance is associated with an Infection Precision of 0.83, with an overall accuracy of 91% (F1 score of 0.91). Subsequently, the sensitivity and specificity analysis revealed that the combination of GZMK, IL7R, mir-93-5p, and mir-345-5p reached good results, with an Area Under the Curve (AUC) of 87% (Figure 1F).

In summary, these analyses showed that, although the power analysis highlighted a limited statistical power due to the small sample size, the model based on the combined marker panel exhibited a promising predictive performance.

4 Discussion

Heart-transplanted patients represent a fragile population. After surgery, they need to be strictly monitored to contrast possible rejection episodes. The primary defense against rejection is immunosuppressive therapy, which patients must take for their entire lives. Unfortunately, these treatments make them more susceptible to bacterial and viral infections. Cytomegalovirus is the most clinically relevant infection in cardiac transplanted patients during the first year of follow-up (1). CMV *de novo* infection or reactivation needs to be treated properly (2, 8). Thus, correct diagnosis is crucial to define the appropriate therapeutic protocol. With this study, we aimed to explore the role of mRNA and miRNA as new biomarkers to improve EMB interpretation and distinguish rejection from infection-related inflammatory infiltrate on graft tissue.

Current molecular tests to assess CMV infection on EMB specimens need improvement. Our study suggests that IL7R and GZMK are decisive in the discrimination between an activate inflammatory status, due to rejection and infection, and control stable patients. These genes were downregulated in the Infection group (-38.63 Fold-change for IL7R and -3.15 Fold-change for GZMK), compared to the Rejection group, suggesting an alteration in the regulatory mechanisms of their pathways. Many groups showed that IL7R (also called CD127) is downregulated in viral infection characterized by a latent phase (38–41). Interestingly,

immune activation and CD4 lymphopenia have been associated with decreased IL7R expression (41). These results perfectly fit with the conditions observed in CMV-positive transplanted patients, who are exposed to a viral-induced stimulation in an impaired immunocompetence condition. Conversely, only a few and contrasting studies have been published about GZMK expression in viral infection (42). Of relevance, Verschoor and colleagues showed that GZMK expression decreased in some specific subsets of T cells in CMV-positive patients, and in older patients, who are known to be characterized by a fragile and altered immune response (43). Focusing on the connection between the pathways of GZMK and IL7R and the miRNA dysregulated in our comparison, we identified two key miRNAs that are up-regulated in the Rejection group: miR-93-5p (5.98 Avg (log2) in Rejection vs 4.59 Avg (log2) in Infection) and miR-345-5p (2.18 Avg (log2) in Rejection vs 1.6 Avg (log2) in Infection). These elements represent an important starting point for investigating new CMV diagnosis biomarkers. We could combine gene expression analysis (IL7R and GZMK) to assess the inflammatory status of the graft, and on the other hand, the assessment of miR-93-5p and miR-345-5p could be helpful to discriminate infection, where their levels of expression are decreased, from rejection. In clinical practice, the possible application of this novel and promising approach could be divided into two phases. First of all, thanks to IL7R and GZMK testing, we could perform a screening process and discriminate between stable and unstable patients, who show IL7R and GZMK increased expression levels compared to non-rejecting patients. Secondly, through mir-93-5p and mir-345-5p analysis, we could improve our capacity to distinguish infection, where those miRNAs show a lower level of expression, from rejection, where they are overexpressed. In this ideal new workflow, combining different biomarkers should be decisive in discriminating between infection and rejection. Indeed, inflammatory response is a dynamic process, and using a double level of investigation could improve our ability to identify infection-associated inflammatory infiltrate on tissue samples.

Over the past 30 years, multiple solid molecular and epidemiological evidences have connected CMV infection and the exacerbation of acute and chronic allograft rejection in solid organ transplants, especially heart transplants (4, 7, 44–47). The association between CMV and CAV, CAD, and vascular dysfunction has been deeply investigated at the molecular level, demonstrating the primary role of CMV in inducing focal inflammation, triggering atherosclerosis, and microvascular damage. These results were also confirmed in human hand transplantation and facial vascularized composite allotransplantation (48–50).

Additionally, this novel paradigm could be effective in improving CMV diagnosis and overcoming the limitations of molecular tests currently in use. Indeed, tests based on the detection of the viral genome are prone to false negative results, due to the focal reactivation of CMV and to sampling error of EMB. The application of PCR or immunohistochemistry analysis to detect the viral genomes or proteins may yield false-negative results when the EMBs fail to sample regions of focal CMV infection. Conversely, PCR may also produce false-positive findings by detecting viral particles present on tissue samples as a result of blood contamination rather than true myocardial viral localization. The analysis of genes that test the immune response to the virus could overcome these limits. Finally, thanks to this *proof-of-concept* study, we can optimistically foresee a wide clinical application of our promising results with cost-effective and user-friendly techniques, such as quantitative PCR (qPCR). Indeed, we identified four molecular candidate markers that could be easily analyzed on EMB specimen extract in all clinical laboratories, with no need for new platforms and minimal equipment costs.

Our study has some limitations. Unfortunately, our enrolled population is quite small, and this is due to the strict inclusion criteria applied in the selection phase. We aimed to evaluate the transcriptomic profile of the EMB in a setting of CMV infection without any other viral or bacterial coinfections, significantly reducing the number of eligible patients. Despite the modest population size, we obtained a huge amount of data thanks to the high-throughput transcriptomic platform employed, without any bias in transcript selection. Our cross-validation results demonstrated good performance, supporting the robustness of our findings despite the relatively low expected values. To address the current limitations, further validation is warranted in an independent cohort, ideally comprising a larger and more heterogeneous patient population. This will enhance the generalizability of our preliminary results to real-world clinical practice and help mitigate the risk of false negatives.

In conclusion, we explored the possibility of identifying new biomarker candidates in post-transplantation infection diagnosis. Our study is a *proof-of-concept* that infection and rejection are characterized by different transcriptomic profiles, both at mRNA and miRNA levels, shedding new light on the understanding of the pathophysiologic mechanisms that underlie these insidious conditions. Combining CMV-related gene analysis with rejection-related miRNA assessment could further improve our "resolution power" in distinguishing CMV infection from ACR, which can show an overlapping phenotype on histology. These novel and interesting results open up new possible innovative, and promising approaches for CMV infection diagnosis with the use of cost-effective qPCR applications, reduced turnaround time, and large-scale clinical applicability of our results.

Data availability statement

The microarray data have been deposited on Zenodo and they are available at this doi: 10.5281/zenodo.15295266. See the link: https://doi.org/10.5281/zenodo.15295266.

Ethics statement

The studies involving humans were approved by Ethics Committee of Azienda Ospedaliera di Padova. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

IB: Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. DP: Data curation, Investigation, Software, Visualization, Writing – review & editing. CC: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. AG: Methodology, Visualization, Writing – review & editing. ER: Methodology, Writing – review & editing. LV: Data curation, Software, Writing – review & editing. SAM: Methodology, Writing – review & editing. CT: Resources, Writing – review & editing. NP: Resources, Writing – review & editing. GT: Resources, Writing – review & editing. FT: Resources, Writing – review & editing. CB: Resources, Writing – review & editing. GG: Resources, Writing – review & editing. SM: Methodology, Resources, Writing – review & editing. DA: Data curation, Supervision, Validation, Writing – review & editing. DG: Data curation, Supervision, Writing – review & editing. AA: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. MF: Conceptualization, Funding acquisition, Investigation, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by University of Padova grants number BIRD191573 and BIRD170215.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1581151/full#supplementary-material

References

1. Potena L, Solidoro P, Patrucco F, Borgese L. Treatment and prevention of cytomegalovirus infection in heart and lung transplantation: an update. *Expert Opin Pharmacother*. (2016) 17:1611–22. doi: 10.1080/14656566.2016.1199684

2. Kotton CN, Kumar D, Caliendo AM, Huprikar S, Chou S, Danziger-Isakov L, et al. The third international consensus guidelines on the management of cytomegalovirus in solid-organ transplantation. *Transplantation*. (2018) 102:900–31. doi: 10.1097/TP.0000000000002191

3. Mendez-Eirin E, Paniagua-Martín MJ, Marzoa-Rivas R, Barge-Caballero E, Grille-Cancela Z, Cañizares A, et al. Cumulative incidence of cytomegalovirus infection and disease after heart transplantation in the last decade: effect of preemptive therapy. *Transplant Proc*. (2012) 44:2660–2. doi: 10.1016/j.transproceed.2012.09.035

4. Saldan A, Mengoli C, Sgarabotto D, Fedrigo M, Angelini A, Feltrin G, et al. Human cytomegalovirus and Epstein–Barr virus infections occurring early after transplantation are risk factors for antibody-mediated rejection in heart transplant recipients. *Front Immunol*. (2023) 14:1171197. doi: 10.3389/fimmu.2023.1171197

5. Valantine HA. The role of viruses in cardiac allograft vasculopathy. *Am J Transpl*. (2003) 4:169–77. doi: 10.1046/j.1600-6143.2003.00316.x

6. Delgado JF, Manito N, Almenar L, Crespo-Leiro M, Roig E, Segovia J, et al. Risk factors associated with cytomegalovirus infection in heart transplant patients: a prospective, epidemiological study. *Transpl Infect Dis*. (2011) 13:136–44. doi: 10.1111/j.1399-3062.2010.00573.x

7. Potena L, Valantine HA. Cytomegalovirus-associated allograft rejection in heart transplant patients. *Curr Opin Infect Dis*. (2007) 20:425–31. doi: 10.1097/QCO.0b013e328259c33b

8. Razonable RR, Humar A. Cytomegalovirus in solid organ transplant recipients—Guidelines of the American Society of Transplantation Infectious Diseases Community of Practice. *Clin Transpl*. (2019) 33:1–23. doi: 10.1111/ctr.13512

9. Kalil AC, Levitsky J, Lyden E, Stoner J, Freifeld AG. Meta-analysis: the efficacy of strategies to prevent organ disease by cytomegalovirus in solid organ transplant recipients. *Ann Intern Med*. (2005) 143:870. doi: 10.7326/0003-4819-143-12-200512200-00005

10. Crespo-Leiro MG, Stypmann J, Schulz U, Zuckermann A, Mohacsi P, Bara C, et al. Clinical usefulness of gene-expression profile to rule out acute rejection after heart transplantation: CARGO II. *Eur Heart J*. (2016) 37:2591–601. doi: 10.1093/eurheartj/ehv682

11. Halloran PF, Madill-Thomsen KS. The molecular microscope diagnostic system: assessment of rejection and injury in heart transplant biopsies. *Transplantation*. (2023) 107:27–44. doi: 10.1097/TP.0000000000004323

12. Agbor-Enoh S, Shah P, Tunc I, Hsu S, Russell S, Feller E, et al. Cell-free DNA to detect heart allograft acute rejection. *Circulation*. (2021) 143:1184–97. doi: 10.1161/CIRCULATIONAHA.120.049098

13. Goldberg JF, Truby LK, Agbor-Enoh S, Jackson AM, deFilippi CR, Khush KK, et al. Selection and interpretation of molecular diagnostics in heart transplantation. *Circulation*. (2023) 148:679–94. doi: 10.1161/CIRCULATIONAHA.123.062847

14. Halloran PF, Potena L, Van Huyen JPD, Bruneval P, Leone O, Kim DH, et al. Building a tissue-based molecular diagnostic system in heart transplant rejection: The heart Molecular Microscope Diagnostic (MMDx) System. *J Heart Lung Transpl*. (2017) 36:1192–200. doi: 10.1016/j.healun.2017.05.029

15. Loupy A, Lefaucheur C, Vernerey D, Chang J, Hidalgo LG, Beuscart T, et al. Molecular microscope strategy to improve risk stratification in early antibody-mediated kidney allograft rejection. *J Am Soc Nephrol*. (2014) 25:2267–77. doi: 10.1681/ASN.2013111149

16. Seddiki N, Santner-nanan B, Martinson J, Zaunders J, Sasson S, Landay A, et al. Expression of interleukin (IL) -2 and IL-7 receptors discriminates between human regulatory and activated T cells. *J Exp Med*. (2010) 203:1693–700. doi: 10.1084/jem.20060468

17. Michel L, Soulillou JP, Laplaud DA, Michel L, Berthelot L, Pettré S, et al. Patients with relapsing-remitting multiple sclerosis have normal Treg function when cells expressing IL-7 receptor α -chain are excluded from the analysis. *J Clin Invest*. (2008) 118:3411–9. doi: 10.1172/JCI35365

18. MacKall CL, Fry TJ, Gress RE. Harnessing the biology of IL-7 for therapeutic application. *Nat Rev Immunol*. (2011) 11:330–42. doi: 10.1038/nri2970

19. Carrette F, Surh CD. IL-7 signaling and CD127 receptor regulation in the control of T cell homeostasis. *Semin Immunol*. (2012) 24:209–17. doi: 10.1016/j.smim.2012.04.010

20. Belarif L, Mary C, Jacquemont L, Mai HL, Danger R, Hervouet J, et al. IL-7 receptor blockade blunts antigen-speci fi c memory T cell responses and chronic in fl ammation in primates. *Nat Commun*. (2018) 9:4483. doi: 10.1038/s41467-018-06804-y

21. Dooms H. Interleukin-7: Fuel for the autoimmune attack. *J Autoimmun*. (2013) 45:40–8. doi: 10.1016/j.jaut.2013.06.007

22. Piotti G, Ma J, Adams E, Cobbold S, Waldmann H. Guiding postablative lymphocyte reconstitution as a route toward transplantation tolerance. *Am J Transpl*. (2014) 14:1678–89. doi: 10.1111/ajt.12756

23. Mai H, Boeffard F, Longis J, Danger R, Martinet B. IL-7 receptor blockade following T cell depletion promotes long-term allograft survival. *J Clin Invest*. (2014) 124:1723–33. doi: 10.1172/JCI66287

24. Bovenschen N, Quadir R, Berg ALVD, Brenkman AB, Vandenberghe I, Devreese B, et al. Granzyme K displays highly restricted substrate specificity that only partially overlaps with granzyme A *. *J Biol Chem*. (2009) 284:3504–12. doi: 10.1074/jbc.M806716200

25. Plasman K, Demol H, Bird PI, Gevaert K, Damme PV. Substrate specificities of the granzyme tryptases A and K. *J Proteome Res*. (2014) 13:6067–77. doi: 10.1021/pr500968d

26. Chowdhury D, Lieberman J. Death by a thousand cuts: granzymes pathways of programmed cell death. *Annu Rev Immunol*. (2008) 26:389–420. doi: 10.1146/annurev.immunol.26.021607.090404

27. Turner CT, Zeglinski MR, Richardson KC, Zhao H, Shen Y, Papp A, et al. Granzyme K expressed by classically activated macrophages contributes to inflammation and impaired remodeling. *J Invest Dermatol*. (2019) 139:930–9. doi: 10.1016/j.jid.2018.09.031

28. Zeglinski MR, Granville DJ. Granzymes in cardiovascular injury and disease. *Cell Signal*. (2020) 76:109804. doi: 10.1016/j.cellsig.2020.109804

29. Poli A, Michel T, Thérésine M, Andrès E, Hentges F, Zimmer J. CD56 bright natural killer (NK) cells: an important NK cell subset. *Immunology*. (2009) 126:458–65. doi: 10.1111/j.1365-2567.2008.03027.x

30. Sharma M, Merkulova Y, Raithatha S, Parkinson LG, Shen Y. Extracellular granzyme K mediates endothelial activation through the cleavage of protease-activated receptor-1. *FEBS J*. (2016) 283:1734–47. doi: 10.1111/febs.2016.283.issue-9

31. Li N, Miao Y, Shan Y, Liu B, Li Y, Zhao L, et al. Mir-106b and mir-93 regulate cell progression by suppression of pten via pi3k/akt pathway in breast cancer. *Cell Death Dis*. (2017) 8:e2796–11. doi: 10.1038/cddis.2017.119

32. Myers MP, Pass I, Batty IH, van der Kaay J, Stolarov JP, Hemmings BA, et al. The lipid phosphatase activity of PTEN is critical for its tumor supressor function. *Proc Natl Acad Sci U S A*. (1998) 95:13513–8. doi: 10.1073/pnas.95.23.13513

33. Tamura M, Gu J, Matsumoto K, Aota S, Parson R, Yamada KM. Inhibition of cell migration, spreading, and focal adhesions by tumor suppressor PTEN. *Science*. (1998) 280:1614–7. doi: 10.1126/science.280.5369.1614

34. Fang L, Du WW, Yang W, Rutnam ZJ, Peng C, Li H, et al. MiR-93 enhances angiogenesis and metastasis by targeting LATS2. *Cell Cycle*. (2012) 11:4352–65. doi: 10.4161/cc.22670

35. Fang L, Deng Z, Shatseva T, Yang J, Peng C, Du WW, et al. MicroRNA miR-93 promotes tumor growth and angiogenesis by targeting integrin-β8. *Oncogene*. (2011) 30:806–21. doi: 10.1038/onc.2010.465

36. Wang P, Fang Y, Qiu J, Zhou Y, Wang Z, Jiang C. miR-345–5p curbs hepatic stellate cell activation and liver fibrosis progression by suppressing hypoxia-inducible factor-1alpha expression. *Toxicol Lett*. (2022) 370:42–52. doi: 10.1016/j.toxlet.2022.09.008

37. Liu J, Jiang Y, Han M, Jiang L, Liang D, Li S, et al. MicroRNA-345-5p acts as an anti-inflammatory regulator in experimental allergic rhinitis via the TLR4/NF-κB pathway. *Int Immunopharmacol*. (2020) 86. doi: 10.1016/j.intimp.2020.106522

38. Crawley AM, Angel JB. Expression of γ-chain cytokine receptors on $CD8^+$ T cells in HIV infection with a focus on IL-7Rα (CD127). *Immunol Cell Biol*. (2012) 90:379–87. doi: 10.1038/icb.2011.66

39. Golden-Mason L, Burton JR, Castelblanco N, Klarquist J, Benlloch S, Wang C, et al. Loss of IL-7 receptor alpha-chain (CD127) expression in acute HCV infection associated with viral persistence. *Hepatology*. (2006) 44:1098–109. doi: 10.1002/hep.21365

40. Boutboul F, Puthier D, Appay V, Pelle O, Ait-Mohand H, Combadière B, et al. Modulation of interleukin-7 receptor expression characterizes differentiation of CD8 T cells specific for HIV, EBV and CMV. *AIDS*. (2005) 19:1981–6. doi: 10.1097/01.aids.0000191919.24185.46

41. Koesters SA, Alimonti JB, Wachihi C, Matu L, Anzala O, Kimani J, et al. IL-7Rα expression on $CD4^+$ T lymphocytes decreases with HIV disease progression and inversely correlates with immune activation. *Eur J Immunol*. (2006) 36:336–44. doi: 10.1002/eji.200535111

42. Bouwman AC, Van Daalen KR, Crnko S, Ten Broeke T, Bovenschen N. Intracellular and extracellular roles of granzyme K. *Front Immunol*. (2021) 12:677707. doi: 10.3389/fimmu.2021.677707

43. Verschoor CP, Picard E, Andrew MK, Haynes L, Loeb M, Pawelec G, et al. NK- and T-cell granzyme B and K expression correlates with age, CMV infection and influenza vaccine-induced antibody titres in older adults. *Front Aging*. (2023) 3:1098200. doi: 10.3389/fragi.2022.1098200

44. Potena L, Holweg CTJ, Chin C, Luikart H, Weisshaar D, Narasimhan B, et al. Acute rejection and cardiac allograft vascular disease is reduced by suppression of subclinical cytomegalovirus infection. *Transplantation*. (2006) 82:398–405. doi: 10.1097/01.tp.0000229039.87735.76

45. Chang A, Musk M, Lavender M, Wrobel J, Yaw M, Lawrence S, et al. Cytomegalovirus viremia in lung transplantation during and after prophylaxis. *Transpl Infect Dis*. (2019) 21:e13069. doi: 10.1111/tid.2019.21.issue-3

46. Korneffel K, Mitro G, Buschor K, Rees M, Ortiz J. Low dose valganciclovir as cytomegalovirus prophylaxis in post-renal transplant recipients induced with alemtuzumab: A single-center study. *Transpl Immunol*. (2019) 56:101226. doi: 10.1016/j.trim.2019.101226

47. Singh N, Winston DJ, Razonable RR, Lyon GM, Silveira FP, Wagener MM, et al. Effect of preemptive therapy vs antiviral prophylaxis on cytomegalovirus disease in seronegative liver transplant recipients with seropositive donors: A randomized clinical trial. *JAMA J Am Med Assoc*. (2020) 323:1378–87. doi: 10.1001/jama.2020.3138

48. Schneeberger S, LucChina S, Lanzetta M, Brandacher G, Bösmüller C, Steurer W, et al. Cytomegalovirus-related complications in human hand transplantation. *Transplantation*. (2005) 80:441–7. doi: 10.1097/01.tp.0000168454.68139.0a

49. Kauke-Navarro M, Panayi AC, Formica R, Marty F, Parikh N, Foroutanjazi S, et al. Cytomegalovirus-related complications and management in facial vascularized composite allotransplantation: an international multicenter retrospective cohort study. *Transplantation*. (2022) 106:2031–43. doi: 10.1097/TP.0000000000004132

50. Huelsboemer L, Boroumand S, Kochen A, Dony A, Moscarelli J, Hauc SC, et al. Immunosuppressive strategies in face and hand transplantation: a comprehensive systematic review of current therapy regimens and outcomes. *Front Transpl*. (2024) 3:1366243. doi: 10.3389/frtra.2024.1366243

ELSEVIER

https://www.jhltopen.org/

The value of systematic study and biobanking of explanted hearts: Insights from an international ISHLT pathology survey[1]

Marny Fedrigo, MD PhD,[a] Diego Perazzolo, Ing,[a] Gerald J. Berry, MD,[b] and Annalisa Angelini, MD[a,*]

[a]*Department of Cardiac, Thoracic, Vascular Sciences, and Public Health, University of Padua, 35128 Padua, Italy*
[b]*Department of Pathology, Stanford University, Stanford, CA*

KEYWORDS:
Heart transplantation;
Explanted heart;
Biobanking;
Pathology;
Survey

This study evaluates global practices for managing explanted hearts, with a focus on tissue collection and biobanking protocols. A survey conducted through the International Society for Heart and Lung Transplantation (ISHLT) assessed responses from centers across Europe, North America, and Other Countries. Results demonstrated significant variability in tissue sampling, grossing protocols, and storage practices. While 78.8% of centers had grossing protocols, fewer (73.1%) adapted sampling based on pathology. Fresh tissue collection was prevalent in 63.5% of centers, but volumes varied: North America led with higher sampling rates (10-25 samples per heart), while Europe and Other Countries collected fewer samples. Coronary artery sampling also showed regional differences. Fresh tissues enable advanced molecular studies, while fixed tissues remain fundamental for histopathology. Standardized global protocols for sampling, storage, and reporting could enhance the clinical and research value of explanted hearts, optimizing post-transplant care and driving innovation in cardiac medicine.

JHLT Open 2025;9:100306

The analysis of explanted hearts offers clinicians a unique opportunity to understand the underlying pathologies leading to transplantation, confirm diagnoses, and contribute to the advancement of cardiac research.[1–10] Histopathological evaluation of formalin-fixed tissues remains indispensable for clinical diagnosis,[5–10] while fresh tissue enables cutting-edge molecular analyses such as transcriptomics, proteomics, and genetic studies.[1–4]

Despite its importance, significant variability persists in how centers collect, store, and analyze explanted hearts. These differences impact both clinical decision-making and the comparability of research data across institutions. In preparation for an International Society for Heart and Lung Transplantation-supported development of protocols for standardized explanted heart evaluation and reporting, a

[1] We report relevant insights from an ISHLT Pathology survey on explanted hearts at transplantation. We highlighted the role of pathologists in biobanking heart tissue and the need of systematic study of the heart removed at transplantation or at autopsy.

*Corresponding author: Annalisa Angelini, MD, Pathology of Cardiac Transplantation and Regenerative Medicine Unit, Department of Cardiac, Thoracic and Vascular Sciences and Public Health, University of Padua, Via A. Gabelli, 61 35121 Padova, Italy. Telephone: +39 049 8272260.

E-mail address: annalisa.angelini@unipd.it.

https://doi.org/10.1016/j.jhlto.2025.100306

Table 1 Heart Storage Practices and Duration by Region

(A) Protocols and practices report Answer	Grossing protocol	X-ray performed	Photographic documentation	Standard form for macroscopy reporting	Blood collection	Different sample according to pathology
Yes (%)	78.85	15.38	78.85	69.23	23.08	73.08
No (%)	21.15	84.62	5.77	30.77	73.08	26.92
Seldom (%)	0.0	0.0	15.38	0.0	3.85	0.0
Europe Yes (%)	75.0	12.5	79.17	66.67	25.0	62.5
Europe No (%)	25.0	87.5	8.33	33.33	70.83	37.5
Europe Seldom (%)	0.0	0.0	12.5	0.0	4.17	0.0
US/Canada Yes (%)	95.24	23.81	76.19	76.19	19.05	80.95
US/Canada No (%)	4.76	76.19	4.76	23.81	76.19	19.05
US/Canada Seldom (%)	0.0	0.0	19.05	0.0	4.76	0.0
Other Countries Yes (%)	42.86	0.0	85.71	57.14	28.57	85.71
Other Countries No (%)	57.14	100.0	0.0	42.86	71.43	14.29
(B) Percentage Number of samples collected from the fresh tissue						
	0 (%)	**1 (%)**	**2 (%)**	**3 (%)**	**4 (%)**	**5 (%)**
Number of samples of fresh tissue per heart	36.54	15.38	23.08	11.54	9.62	3.85
(C) Percentage Number of samples collected from fixed tissue						
	1-10 (%)	**10-25 (%)**	**> 25 (%)**			
How many fixed samples do you take for each heart?	55.32	38.3	6.38			
(D) Percentage Number of samples collected from the coronary arteries						
	0 (%)	**1 (%)**	**2 (%)**	**3 (%)**	**4 (%)**	**> 4 (%)**
How many samples do you take for the coronary arteries?	4.35	8.7	21.74	36.96	15.22	13.04

Panel A: The table highlights differences across regions in the adoption of multiple practices showing areas where certain protocols are more commonly implemented. Panel B: This section shows that most centers collect one or no samples of fresh tissue per heart, with fewer centers collecting larger numbers, indicating varied tissue collection practices. Panel C: The data indicates that the majority of centers collect between 1 and 10 fixed tissue samples per heart. Multiple centers collect a number between 10 and 25, while only few of them collect a number of samples higher than 25. Panel D: This section demonstrates that centers typically collect 1 to 3 coronary artery samples, with fewer centers collecting larger numbers. Only few of the centers collected 0 samples.

survey was organized to characterize current transplant center practices in tissue collection, storage, and examination of explanted hearts.

This study highlights key findings, including grossing protocols, sampling practices for fresh and fixed tissues, and storage durations. By promoting standardized protocols, clinicians and researchers can maximize the diagnostic and research value of these precious specimens.

Methods

A structured survey was distributed to International Society for Heart and Lung Transplantation-affiliated heart transplant centers across three global regions: Europe, North America (United States/Canada), and Other Countries. The survey explored the following areas:

1. Grossing and Tissue Sampling Protocols:

- o Presence of grossing protocols.
- o Collection of fresh and formalin-fixed tissue.
- o Number of samples collected per heart (fresh and fixed).
- o Coronary artery sampling practices.

2. Storage Practices:

- o Conditions for short- and long-term storage (e.g., frozen, formalin-fixed).
- o Duration of heart specimen storage.

3. Reporting and Blood Collection:

- o Use of standardized reporting forms.
- o Frequency of blood collection at transplantation.

Quantitative responses were analyzed and expressed as percentages. Descriptive statistics highlighted regional trends, with results summarized in Table 1 and Figures 1, 2.

Results

Grossing Protocols and Reporting Standards: A majority of centers (78.8%) reported having a formal grossing protocol

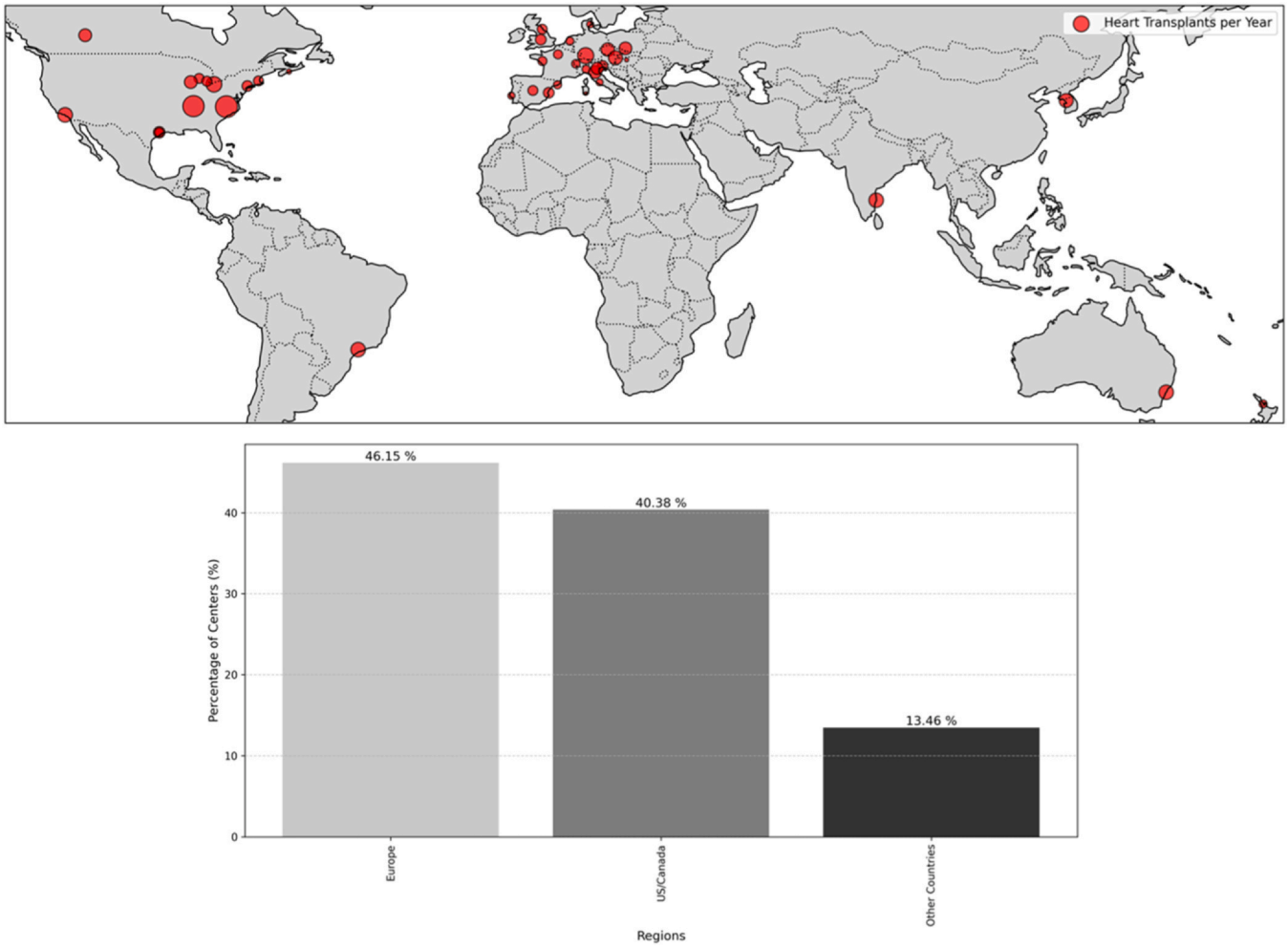


Figure 1 *Distribution of Cardiac Transplant Centers by Region which have answered to the survey.* The figure illustrates the geographic distribution of cardiac transplant centers that responded to the survey, emphasizing regional variations in participation and practices.

(Table 1A), though adoption was highest in North America (95.2%) compared to Europe (75%) and Other Countries (42.9%). The use of standardized reporting forms was similarly prevalent in North America (76.2%) but less so in Other Countries (57.1%).

Fresh and Fixed Tissue Sampling Practices: Fresh tissue sampling was reported by 63.5% of centers globally (Table 1B). North American centers collected the highest number of fresh tissue samples per heart, with 38.3% reporting 10-25 samples, while Europe predominantly collected 1-10 samples (55.3%). In Other Countries, fresh tissue sampling was inconsistent, with most centers collecting fewer than five samples per heart. For fixed tissues, 55.3% of centers reported collecting 1-10 samples, while 38.3% collected 10-25 samples (Table 1C). Coronary artery sampling practices also varied: 36.9% of centers sampled three arteries, and 15.2% sampled four (Table 1D). Storage Practices: Heart storage methods and duration exhibited regional differences (Figure 2A, B). In Europe, the prevailing trend is to store hearts removed during transplantation in nearly two-thirds of cases. In North America, there was a slight decline, with the trend decreasing to approximately to two-fifty of cases. Globally, formalin fixation was the most commonly used method, although some centers reported using freezing methods, particularly when prioritizing molecular research. Storage duration was relatively consistent across Europe, North America, and other regions, with approximately half of the centers storing hearts indefinitely. Notably, only 3 out of 52 centers collected fresh materials outside the pathology department. In all other cases, pathologists were responsible for storing both fresh and fixed materials for biobanking purposes.

Discussion

The findings highlight substantial global variability in the management of explanted hearts, particularly in grossing protocols, tissue sampling, and storage practices. For clinicians, these results underscore the complementary value of fresh and fixed tissue samples. Fresh tissues facilitate molecular and genetic studies that can uncover novel biomarkers and therapeutic targets,[1–4] while formalin-fixed tissues remain the cornerstone for histopathology, confirming diagnoses such as myocarditis, cardiomyopathy, or transplant rejection.[5–10] However, the survey revealed disparities in sample volumes and coronary artery sampling.

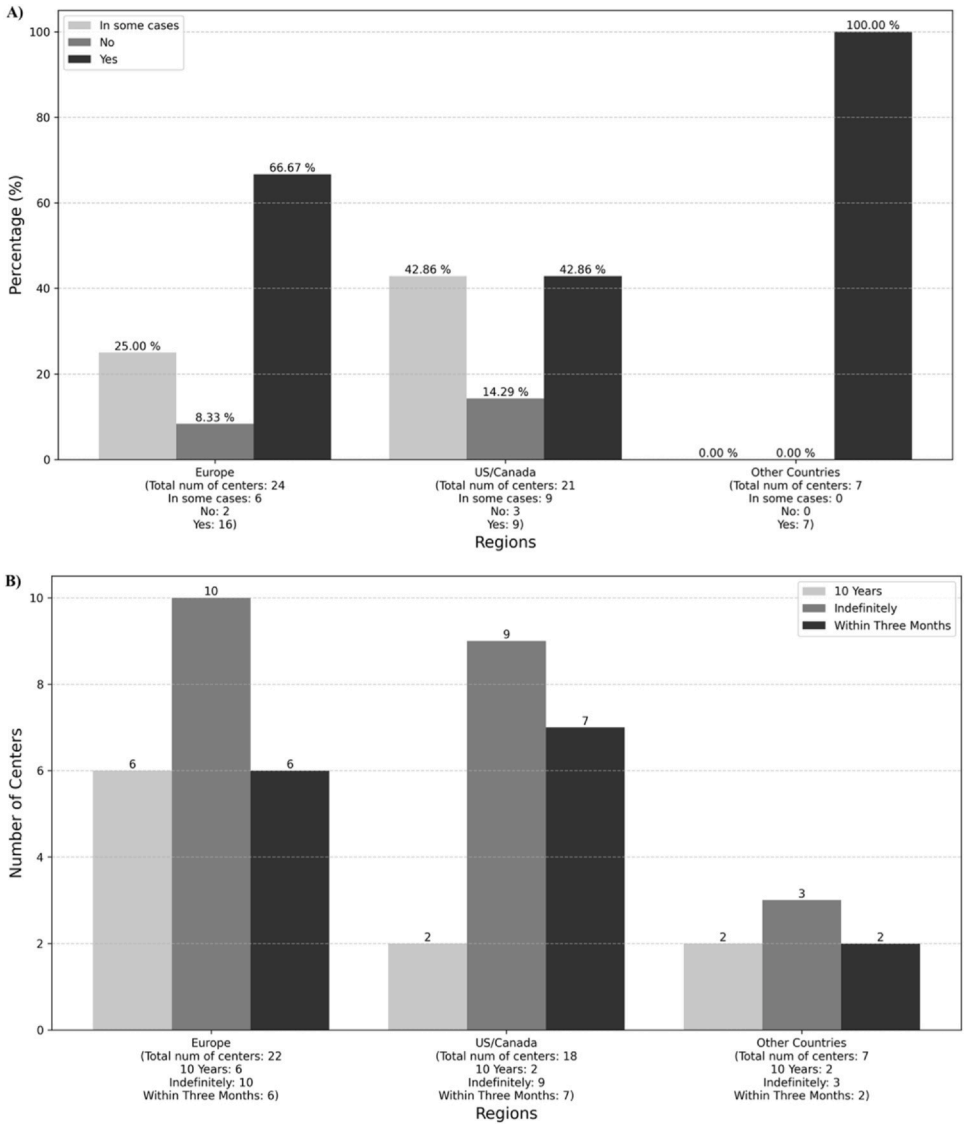


Figure 2 **(A)** *Heart storage practice by region.* This figure shows regional variation in heart storage methods, with some regions favoring specific practices more consistently than others. **(B)** *Heart storage duration distribution across the centers of the region.* The grouped bar plot in this picture illustrates heart storage durations across centers that reported performing storage practices in different regions. It is evident that most centers store hearts for an indefinite period.

North American centers demonstrated higher tissue collection rates, likely reflecting greater resource availability and research integration. In contrast, centers in Other Countries reported limitations due to infrastructure and personnel shortages. Standardized protocols for tissue sampling and storage are urgently needed. Uniform practices, including guidelines for sample volume, chamber selection, and coronary artery sampling, would ensure more consistent clinical evaluations and enhance the comparability of research findings across regions. Additionally, expanding access to freezing and cryopreservation technologies could improve fresh tissue collection in resource-limited settings. The widespread adoption of grossing protocols (78.8%) and standardized reporting forms (69.2%) in North America and Europe highlights their value in improving diagnostic uniformity and accuracy. Training programs aimed at enhancing these practices globally could further bridge the existing gaps.

Conclusion

The systematic collection of fresh and fixed tissues from explanted hearts offers clinicians an invaluable resource for both diagnostics and research. This survey identifies key areas of variability in tissue sampling, grossing protocols, and storage practices across global transplant centers. Promoting standardized protocols will optimize the clinical and scientific value of explanted hearts, supporting more precise post-transplant care and driving innovations in cardiac medicine.

CRediT authorship contribution statement

AA and MF formed idea for the survey and this manuscript. DP made the statistical analysis of the survey and designed the illustrations. GB contribute to the data revision. AA wrote the manuscript which has been read and revised by all the authors.

Disclosure statement

The authors of this manuscript have no conflicts of interest to disclose, as described by the Journal of heart and lung transplantation. No funding was received for this study.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

For data ollection and statistical data analysis none artificial intelligence (AI) application have been used. AI based application have been involved to assist the writing procedure of the text.

Acknowledgments

The authors would like to thank all the pathologists who have performed the survey for data collection used in this study. The list of disclosed and undisclosed pathologists and centers are listed in Supplementary Materials.

Appendix A. Supporting Information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jhlto.2025.100306.

References

1. Coppola L, Cianflone A, Grimaldi AM, et al. Biobanking in health care: evolution and future directions. J Transl Med 2019;17:172. https://doi.org/10.1186/s12967-019-1922-3. PMID: 31118074. PMCID: PMC6532145.
2. Malsagova K, Kopylov A, Stepanov A, et al. Biobanks-A platform for scientific and biomedical research. Diagnostics 2020;10:485. https://doi.org/10.3390/diagnostics10070485. PMID: 32708805.PMCID: PMC7400532.
3. Zhu Y, Jackson D, Hunter B, et al. Models of cardiovascular surgery biobanking to facilitate translational research and precision medicine. ESC Heart Fail 2022;9:21-30. https://doi.org/10.1002/ehf2.13768. PMID: 34931483. PMCID: PMC8787984.
4. Song J, Xing Y, Chen X, et al. Processing of the explanted heart. North Am J Med Sci 2014;6:613-7. https://doi.org/10.4103/1947-2714.147975. PMID: 25599048. PMCID: PMC4290049.
5. Ottaviani G, Radovancevic R, Kar B, Gregoric I, Buja LM. Pathological assessment of end-stage heart failure in explanted hearts in correlation with hemodynamics in patients undergoing orthotopic heart transplantation. Cardiovasc Pathology 2015;24:283-9. https://doi.org/10.1016/j.carpath.2015.06.002. 617.
6. Zhang H, et al. The human explanted heart program: a translational bridge for cardiovascular medicine. Biochim Biophys Acta Mol Basis Dis 2020;1867:165995. https://doi.org/10.1016/j.bbadis.2020.165995. PMCID: PMC7581399 PMID: 33141063.
7. Lal S, Li A, Allen D, et al. Best practice bio banking of human heart tissue. Biophys Rev 2015;7:399-406. https://doi.org/10.1007/s12551-015-0182-6. PMCID: PMC4792521 NIHMSID: NIHMS735099 PMID: 26998172.
8. Moreira JD, Gower AC, Xue L, et al. Systematic dissection, preservation, and multiomics in whole human and bovine hearts. Cardiovasc Pathol 2023;63:107495. https://doi.org/10.1016/j.carpath.2022.107495. PMCID: PMC10031913 NIHMSID: NIHMS1873573 PMID: 36334690.
9. Leone O, Angelini A, Bruneval P, Potena L, editors. The pathology of cardiac transplantation: a clinical and pathological perspectives Berlin, Germany: Springer; 2016. https://doi.org/10.1007/978-3-319-46386-5.
10. Michel R, Berry GJ. EDs pathology of transplantation: a practical diagnostic approach. Berlin, Germany: Springer,; 2016. https://doi.org/10.1007/978-3-319-29683-8.