Heliyon 10 (2024) e31247

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

An integrated signature of clinical metrics and immune-related genes as a prognostic indicator for ST-segment elevation myocardial infarction patient survival

Wei Gao [a,b,c,d,**], Xiao-yan Wang [e,1], Xing-jie Wang [f], Lei Huang [a,b,c,d,*]

[a] Department of Heart Center, Tianjin Third Central Hospital, Tianjin, 300170, PR China
[b] Tianjin Key Laboratory of Extracorporeal Life Support for Critical Diseases, Tianjin Third Central Hospital, Tianjin, 300170, PR China
[c] Artificial Cell Engineering Technology Research Center, Tianjin, 300170, PR China
[d] Tianjin Institute of Hepatobiliary Disease, Tianjin, 300170, PR China
[e] Institute of Biomedical Science, Fudan University, Shanghai, 200030, PR China
[f] Clinical Laboratory of Tianjin Chest Hospital, Tianjin, 300222, PR China

ARTICLE INFO

Keywords:
Acute coronary syndrome
S100A12 protein
IL2RB protein
Prognosis
Transcriptome
Bioinformatics

ABSTRACT

Background: The immune-inflammatory pathway plays a critical role in myocardial infarction development. However, few studies have systematically explored immune-related genes in relation to myocardial infarction prognosis using bioinformatic analysis. Our study aims to identify differentially expressed immune-related genes(DEIRGs) in ST-segment elevation myocardial infarction (STEMI) patients and investigate their association with clinical outcomes.

Materials and methods: We conducted a systematic review of Gene Expression Omnibus datasets, selecting GSE49925, GSE60993, and GSE61144 for analysis. DEIRGs were identified using GEO2R and overlapped across the chosen datasets. Functional enrichment analysis elucidated the DEIRGs' biological functions and pathways. We established an optimal prognostic prediction model using LASSO penalized Cox proportional hazards regression. The signature's clinical utility was evaluated through survival analysis, ROC curve assessment, and decision curve analysis. Additionally, we constructed a prognostic nomogram for survival rate prediction. External validation was performed using our own plasma samples.

Results: The resulting prognostic signature integrated two dysregulated DEIRGs (*S100A12* and *IL2RB*) and two clinical variables (serum creatinine level and Gensini score). This signature effectively stratified patients into low- and high-risk groups. Survival analysis, ROC curve analysis, and decision curve analysis demonstrated its robust predictive performance and clinical utility within the first two years post-disease onset. External validation confirmed significant outcome differences between risk groups.

Conclusions: Our study establishes a prognostic signature that combines DEIRGs and clinical variables for STEMI patients. The signature exhibits promising predictive capabilities for patient stratification and survival risk assessment.

* Corresponding author. Department of Heart Center, Tianjin Third Central Hospital, Tianjin, 300170, PR China.
** Corresponding author. Department of Heart Center, Tianjin Third Central Hospital, Tianjin, 300170, PR China.
E-mail address: huangleidoctor@aliyun.com (L. Huang).
[1] These authors contributed equally to this work and share first authorship.

https://doi.org/10.1016/j.heliyon.2024.e31247
Received 14 January 2024; Received in revised form 26 April 2024; Accepted 13 May 2024
Available online 15 May 2024

1. Introduction

Acute myocardial infarction(AMI) constitutes a clinical event characterized by myocardial ischemia leading to tissue injury. Diagnosis of AMI relies on troponin elevation, preferably detected using high-sensitivity assays, alongside corroborative evidence such as typical symptoms, electrocardiographic changes, or imaging findings indicative of new myocardial loss or regional wall motion abnormalities [1].

ST-segment elevation myocardial infarction(STEMI), a critical emergency, necessitates immediate medical attention. Its pathophysiological process involves a cascade of events triggered by abrupt coronary artery occlusion, often due to a ruptured atherosclerotic plaque followed by thrombus formation [2]. The urgent problems in STEMI management include timely diagnosis, effective treatment to minimize myocardial damage, and accurate prognosis to guide long-term care [3]. Despite advances in reperfusion therapies and pharmacological interventions, post-STEMI complications, including heart failure, remain significant concerns. Developing robust prognostic models is crucial for identifying patients at risk of adverse outcomes and tailoring individualized treatment plans [4,5].

Increasingly, innate immunity's role, particularly that of immune-related genes(IRGs), has gained recognition in STEMI onset and progression [6]. Innate immune cells and their cytokines not only contribute to the immediate response to myocardial injury but also influence subsequent healing and remodeling processes [7]. Recent studies have highlighted IRGs as potential biomarkers for early AMI diagnosis and outcome prediction [8–10]. However, a knowledge gap persists regarding the intricate relationship between these IRGs and clinical outcomes in STEMI patients.

Our prior research has illuminated the prognostic significance of specific IRGs, including *CD8A*, *IL2RB*, and *S100A12*, within the context of STEMI [11]. These findings underscore the potential of integrating clinical parameters with IRG profiles to construct a robust predictive model for patient outcomes. Yet, comprehensive studies systematically exploring this integration and its practical applicability in clinical settings remain scarce.

This study aims to bridge this gap by employing systematic bioinformatics methods to identify differentially expressed IRGs (DEIRGs) and correlate them with clinical factors influencing STEMI outcomes. Our goal is to construct a predictive model that not only enhances STEMI prognosis accuracy but also provides insights into underlying mechanisms. Such a model could revolutionize STEMI management, offering personalized treatment and care.

2. Materials and Methods

2.1. Selection of the expression profile dataset

We meticulously curated datasets for further evaluation based on the following stringent requirements: (i) we exclusively considered datasets that provided comprehensive mRNA expression data related to STEMI from whole peripheral blood samples; (ii) we focused on discovery datasets containing mRNA expression profiles from STEMI patients who had undergone primary percutaneous coronary intervention(PCI), along with samples from healthy controls. We excluded datasets that met any of the following criteria:(i) Utilized RNA sequencing for detection instead of microarray technology; (ii)Included patients who did not receive primary PCI treatment; and (iii) contained data related to cell types (such as circulating endothelial cells or nucleated cells) not present in plasma. From the 18 datasets initially obtained, we identified six datasets derived from comprehensive transcriptome analyses of plasma. These datasets—namely GSE29111, GSE60993, GSE61144, GSE34198, GSE49925, and GSE34571—represented the most prevalent sample types. However, due to specific limitations [11], we excluded datasets GSE29111, GSE34198, and GSE34571. To facilitate our analysis, we selected three datasets (GSE49925 [12], GSE60993 [13], and GSE61144 [13]) for further investigation.

In our pursuit of identifying DEIRGs that distinguish STEMI patients from healthy individuals, we carefully selected two compact datasets for our initial training phase: (i) GSE60993: This dataset comprises seven samples each from healthy controls and STEMI subjects; (ii) GSE61144: It encompasses ten healthy control samples juxtaposed with seven samples from STEMI patients. For validation purposes, we turned to GSE49925, which includes comprehensive medical records and follow-up data. This dataset facilitated a comparative analysis involving 61 individuals with STEMI against a control group of 93 healthy subjects.

2.2. Identification of DEIRGs

Our primary objective was to identify differentially expressed genes (DEGs) between healthy control patients and those with STEMI. For this purpose, we processed raw microarray data from the GEO database for datasets GSE61144 and GSE60993 using the online tool GEO2R2 (based on the R package "limma"). DEGs were screened based on the criteria of $|\log2(\text{Fold Change})| > 1$ and a false discovery rate (FDR) < 0.05. Subsequently, we extracted DEIRGs from the DEGs. Specifically, we obtained 1793 unique immune-related genes from the Immunology Database and Analysis Portal (ImmPort) database (https://www.immport.org/home). We compared the lists of deregulated genes from the two selected GEO datasets to identify overlapping DEIRGs. Visualization included a Venn diagram created using the R packages "ggplot2" and "VennDiagram," as well as volcano plots and difference ranking plots to characterize DEIRG expression levels. Spearman correlation analysis and heatmap visualization were used to explore expression correlations among DEIRGs in patients with varying prognoses. All analyses were conducted using the R package "ggplot2".

2.3. Functional and pathway enrichment analysis

To gain insights into the biological functions of these DEIRGs, we performed functional enrichment analysis, including Gene Ontology(GO) and Kyoto Encyclopedia of Genes and Genomes(KEGG) analyses using the R package "clusterProfiler." Enrichment was based on an FDR threshold of <0.05. The top three significant GO terms and KEGG signaling pathways were visually represented using R packages such as "ggplot2," "igraph," and "ggraph."

2.4. Construction of A signature integrating DEIRGs

We crafted a prognostic model that integrates DEIRGs with clinical predictors using the GSE49925 dataset. This model was developed through the application of the Least Absolute Shrinkage and Selection Operator(LASSO) method, which was implemented using the Cox proportional hazards regression framework provided by the R packages "survival" and "glmnet". Prognostic clinical factors were pinpointed using a significance threshold of $P < 0.1$. Consequently, we calculate the risk score for each STEMI patient by multiplying the expression level of each parameter by its corresponding coefficient and then summing up the results. Patients within the GSE49925 dataset were then stratified into low-risk and high-risk categories based on the median risk score.

2.5. Clinical utility evaluation of the signature

To evaluate the performance of our prognostic model, we employed a risk score distribution plot alongside a Kaplan–Meier survival curve, utilizing the R package "survival" for the survival analysis of the two risk groups. Further, we conducted a time-dependent receiver operating characteristicROC) curve analysis, encompassing 1-, 2-, and 3-year survival predictions, to gauge the model's sensitivity and specificity, using the R package "survivalROC". Decision curve analysis (DCA) was also utilized to ascertain the model's net benefit. This analytical segment and the visualization of results were facilitated by the R packages "survival" and "stdca.R".

2.6. Prognostic nomogram formulation

To quantitatively predict the survival risk of STEMI patients, we constructed a nomogram based on our established model. Each clinical parameter within the nomogram was assigned a score, and the total score was derived by summing the individual scores of all parameters. Calibration curves were then plotted to compare the predicted survival against the actual outcomes, thus assessing the nomogram's predictive accuracy. The plotting of the nomogram and calibration curves was executed using the R package "rms".

2.7. External validation

92 patients with STEMI admitted to our centers were used as an external validation cohort for this model. The inclusion criteria were: 1) age $\geq$18 years; 2) ESC guidelines-based STEMI diagnosis; 3) symptom onset within 12 h prior to PCI; and 4) informed consent obtained. We excluded patients who presented with any of the following conditions: 1) cardiogenic shock; 2) prior myocardial infarction or coronary artery bypass grafting; 3) severe hepatic or renal dysfunction; 4) cancer or other terminal illnesses; and 5) antiplatelet or anticoagulant therapy intolerance. We collected blood samples from the antecubital vein before PCI and stored them in EDTA tubes at -80 °C for analysis. We extracted total RNA from plasma samples using the miRNeasy Mini Kit (Qiagen, Germany) according to the manufacturer's protocol. We performed reverse transcription using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA) and random primers. We calculated the relative expression levels of the target genes using the 2^-ΔΔCt method, where $\Delta Ct = Ct(\text{target gene}) - Ct(GAPDH)$ and $\Delta\Delta Ct = \Delta Ct(\text{sample}) - \Delta Ct(\text{calibrator})$. The calibrator was a pooled sample of plasma RNA from healthy controls. Patients were divided into high-risk and low-risk groups based on our median risk score value. We followed up the patients prospectively to record major adverse cardiovascular events (MACEs) for a median duration of 345 days, comprising cardiac death, myocardial infarction recurrence, heart failure hospitalization, and revascularization repetition. Kaplan-Meier survival curves and risk distribution plots were employed to evaluate the prognostic value of the model. The study protocol adhered to the Declaration of Helsinki and obtained approval from Tianjin Chest hospital's ethics committee.

2.8. Statistical analysis

SPSS version 24.0 (SPSS Inc., Chicago, IL, United States) and R 4.2.1 (found at http://www.R-project.org, The R Foundation) were employed in this study. Normal distributed quantitative data were given as mean standard deviation and contrasted using *t*-tests or *t'* tests for two-group comparisons in order to characterize and compare the baseline demographic, clinical, and laboratory parameters. Quantitative data with an abnormal distribution were reported as median and interquartile ranges, and Mann-Whitney *U*-tests were used to compare results. Qualitative data were given as frequency and composition, and Fisher's exact test was used to analyze the differences in constituent ratios between two groups. All *P* values were two-sided, and $P < 0.05$ was considered for determining statistical significance.

3. Results

3.1. Identification of DEIRGs

The analysis of 14 STEMI samples and 17 normal samples pooled from GSE60993 and GSE61144 gave rise to 86 DEGs. Afterwards, 17 DEIRGs were extracted from the DEGs (Fig. 1A), including 14 up-regulated and 3 down-regulated genes (Table 1). As shown in the volcano plot (Fig. 1B) and difference ranking plot (Fig. 1C), *MMP9* and *GNLY* are the top up-regulated and down-regulated DEIRGs in the early stage of STEMI patients.

The term that exhibited the highest enrichment across biological processes, cellular components, and molecular functions was "positive regulation of *NF-kappa B* transcription factor activity", "plasma membrane signaling receptor complex", and "immune receptor activity", respectively (Fig. 2A, Supplementary Table 1). These functions are closely related to immune regulation, inflammatory response, vascular permeability, etc. Functional enrichment analysis indicated that the "cytokine–cytokine receptor interaction pathway" was the most relevant to the DEIRGs. These pathways are closely related to myocardial ischemia-reperfusion injury, myocardial remodeling, heart failure, etc. The correspondence between DEIRGs and GO term was visualized via a network in Fig. 2B.

3.2. Construction of an immune-related gene prognostic signature

The prognosis related DEIRGs and clinical parameters were identified, and the prediction signature was established based on the GSE49925 dataset. Based on the Meta data available, the study subjects were observed over an average duration of 2.4 years for the incidence of cardiovascular mortality. Out of the 61 individuals diagnosed with STEMI, 55 remained alive, whereas six succumbed to cardiac-related mortality during the observation period. The demographics and clinical characteristics of the STEMI cohort can be referred to our previously published literature [5].

A principal component analysis(PCA) framework was constructed for the DEIRGs to categorize STEMI patients based on their prognostic outcomes within the dataset. The graphical representation of the initial two principal components revealed a clear demarcation, accounting for a cumulative 65.1 % of the dataset's variance (Supplementary Fig. 1). The expression correlation heatmap intuitively demonstrated that the expression profiles of the DEIRGs in STEMI patients with different prognosis were significantly different (Fig. 3A). Moreover, in the group that experienced mortality, the expression correlation among the DEIRGs was markedly diminished when contrasted with the group that survived (Fig. 3B).

Comparing the DEIRGs in patients with varying prognoses revealed significant upregulation of transcription levels for *S100A12* and *IL2RB* in the mortality group, as opposed to the survival group (Fig. 4A). Next, all available clinical indicators (body mass index, systolic/diastolic blood pressure at admission, white blood cell count, Gensini score, high and low-density lipoprotein cholesterol, triglyceride, total cholesterol, fasting blood glucose, serum creatinine) were introduced into a LASSO penalized Cox proportional hazards regression. Serum creatinine level at admission and Gensini score were identified as candidates of the model. Multivariate Cox regression confirmed that the two clinical parameters were independent prognostic factors for STEMI patients (Table 2).

Here, we conduct a proportional risk hypothesis test on the included variables and the multivariate model to verify whether Cox regression meets the prerequisites for application. The results showed that all the included univariates and the model met $P > 0.05$, indicating that the multivariate model meets the proportional risk assumption (Supplementary Table 2). The biomarker signature was thus formulated using the expression data for the primary indicators, each weighted by their respective Cox regression coefficients, in the following manner: [Expression level of *S100A12* × 0.835]+[Expression level of *IL2RB* × (−1.432)]+[Expression level of Gensini × 0.0201]+[Expression level of creatinine (mg/dl) × 0.297]+1.267.

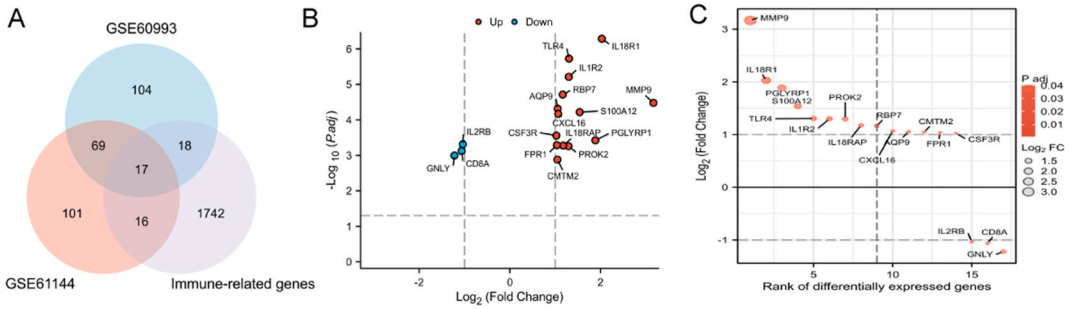

Fig. 1. Identification of DEIRGs from GSE61144 and GSE60993 microarray. (A) Venn diagram of 17 DEIRGs from the two microarray datasets. (B) The volcano plots of DEIRGs in the two datasets. Red indicates genes with high levels of expression, blue indicates genes with low levels of expression based on the criteria of $P < 0.05$ and $|\log(\text{Fold Change})| > 1.0$, respectively. (C) Ranking of the transcriptional difference among the DEIRGs. DEGs differential expression genes, DEIRGs differential expression of immune-related genes.

Table 1
17 DEIRGs were identified from GSE60993, GSE61144 microarrays for STEMI.

DEGs	Gene symbol *
Downregulated (3)	*CD8A, IL2RB, GNLY*
Upregulated (14)	*TLR4, AQP9, CMTM2,CSF3R, PGLYRP1,PROK2, IL1R2, S100A12, RBP7, IL18R1, CXCL16, IL18RAP, MMP9, FPR1*

DEGs, differentially expressed genes; STEMI ST-segment elevated myocardial infarction.

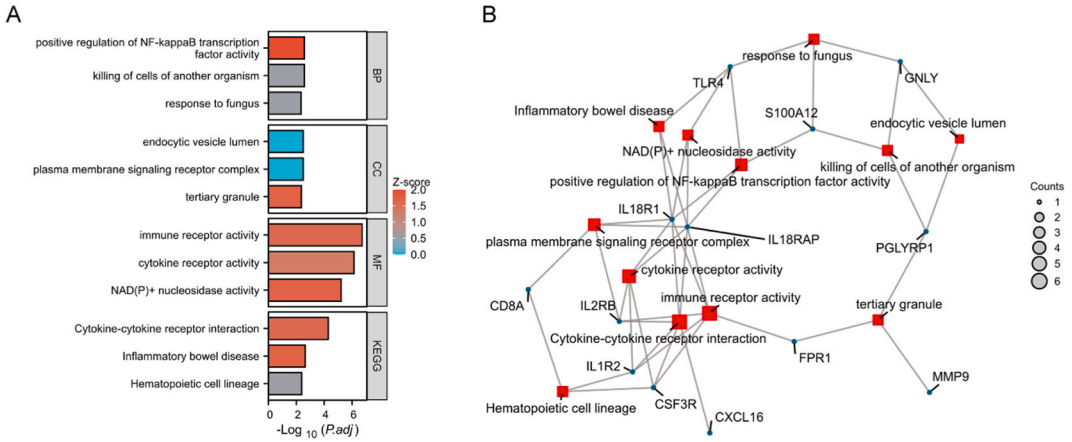

Fig. 2. GO enrichment and KEGG pathway analysis of DEIRGs. (A) Group display of biological processes, cellular components, and molecular functions of DEIRGs. (B) Network of the association between DEIRGs and GO terms and KEGG pathways. GO Gene Ontology, KEGG: Kyoto Encyclopedia of Genes and Genomes, NAD + nicotinamide adenine dinucleotide, ADP adenosine diphosphate, BP for biological process, CC for cellular component and MF for molecular function.

3.3. The signature predicts survival of STEMI patients

For internal validation, the cohort of 61 STEMI patients from the GSE49925 dataset was stratified into groups of lower and higher risk using the median risk score value, with the model's threshold set at −0.418. Illustrated in Fig. 4B, the risk scores associated with the deceased were notably elevated compared to those who survived. A marked disparity in survival rates was evident, with high-risk individuals faring significantly worse than their low-risk counterparts, as depicted in Fig. 4C.

The model's predictive accuracy was gauged using time-dependent ROC curves, showcased in Fig. 5A. The areas under the curve for 1-year, 2-year, and 3-year survival were 0.991, 0.983, and 0.717, respectively, demonstrating the model's robust capability in tracking short-to medium-term survival outcomes. The risk score's cutoff point at the 1-year mark was established at 2.99, and the sensitivity and specificity were 100 % and 96 %, respectively. At 2-year and 3-year follow-up, the above values were 1.467, 100 %, 92 % and 1.467, 49 %, 100 %, respectively. A Kaplan-Meier curve also confirmed that the survival rate of patients in the low-risk group was significantly higher than that in the high-risk group during the 2-year follow-up period (HR = 5.56, Log-rank $P = 0.077$) (Fig. 5B). The results of DCA analysis also showed that the signature had greater clinical net benefits than its component(s) within the first two years after the onset of STEMI (Fig. 5C&D), and this benefit weakened thereafter (Supplementary Fig. 2).

Utilizing the established signature, we developed a prognostic nomogram designed as a quantitative tool for forecasting individual patient survival risks (Fig. 6A). A prognosis calibration was then perform to analyze the fitting between the model established by the Cox regression method and the actual situation, As shown in Fig. 6B, the nomogram's calibration curves demonstrated commendable alignment between the predicted and observed 1- and 2-year survival rates within the STEMI cohort, with a Concordance index (C-index) of 0.936 (ranging from 0.912 to 0.959). However, it was noted that the predictive precision notably declined in the third year.

3.4. The signature is an independent prognostic indicator

To establish the independence of the signature, we conducted multivariate Cox proportional hazards regression analysis within the STEMI cohort from GSE49925. The findings revealed a significant association between the risk score and survival ($P < 0.05$). Consequently, the index was deemed an independent prognostic indicator in the STEMI cohort [HR (95 % CI) = 2.535 (1.124–5.714), $P = 0.025$].

3.5. External validation by our cohorts

We evaluated the expression levels of the two genes in the patients' blood and calculated a risk score using our signature. Based on this score, we categorized patients into low-risk and high-risk groups, using the median as the threshold. The groups were similar in

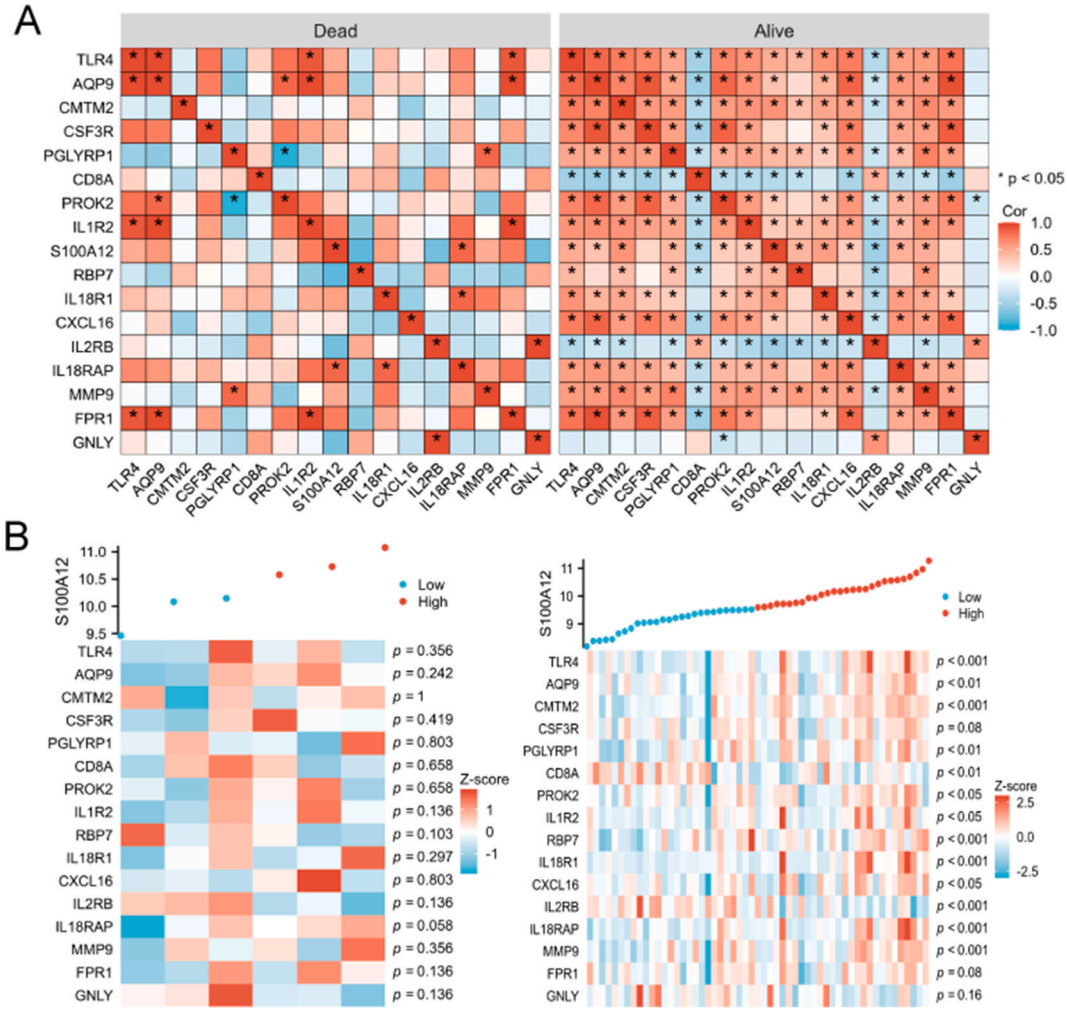

Fig. 3. (A) Expression correlation heatmap of DEIRGs between different prognostic groups of GSE49925. Cor correlation coefficient. (B) Co-expression heat maps between DEIRGs based on an individual level. The correlation was significantly weakened in the death group (left) compared to the survival group (right).

terms of age, gender, smoking status, comorbidities, heart rate, and blood pressure. However, the expression of two genes, as well as serum creatinine and Gensini score, were markedly different between the groups (Table 3). During the follow-up period (median 426 days), the low-risk group had only 2 events, one repeat heart attack and one hospitalization for heart failure; the high-risk group had 8 events, including one death, one repeat heart attack, five hospitalizations for heart failure, and one additional revascularization. The group identified as high-risk exhibited a markedly elevated incidence of events compared to the low-risk group ($P = 0.044$). Illustrated in Fig. 7A, the incidence of actual MACEs was substantially greater in the high-risk group relative to the low-risk group. Fig. 7B presents the Kaplan-Meier survival curves, demonstrating a less favorable survival outcome for the high-risk group (HR = 4.34), with the survival disparity approaching statistical significance (Log-rank $P = 0.068$).

4. Discussion

STEMI is a common cardiovascular disease in emergency and critical cases, whose pathogenesis involves the necrosis of cardiomyocytes and the activation of immune-inflammatory pathways [7,14,15]. In this study, we found that *MMP9* and *GNLY*, as two immune-related genes with the most significant differences in expression, showed consistent trends of expression changes in early serum of STEMI patients with previous research conclusions [16,17]. Matrix metalloproteinases(MMPs) constitute a group of enzymes responsible for extracellular matrix degradation, significantly impacting myocardial remodeling [18]. Specifically, MMP9 plays a role in the inflammatory response, fibrosis, and angiogenesis following AMI [18]. Previous research has consistently demonstrated elevated MMP9 levels in the serum of AMI patients, correlating with an unfavorable prognosis [17,19]. The upregulation of MMP9 may be due to the stimulation of factors such as myocardial cell necrosis, inflammatory cell infiltration, oxidative stress, etc. caused by myocardial ischemia [16]. The high expression of *MMP9* may lead to plaque instability, increased fragility, fibrinolysis inhibition, etc., increasing

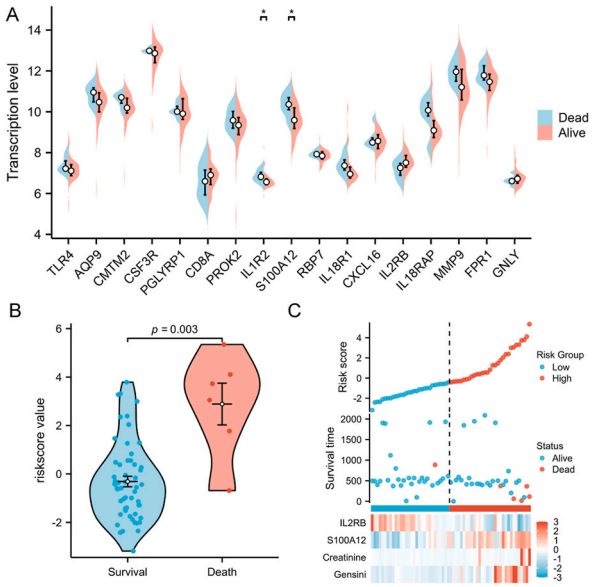

Fig. 4. GSE49925 validated the survival prediction of this signature. (A) Comparison of 17 DEIRGs in patients with different prognosis of GSE49925 (Wilcoxon rank sum test, $^*P < 0.05$). (B) Comparison of risk scores between different prognostic groups. (C) A risk distribution plot intuitively shows the significant difference of survival outcomes between the two groups with different risk stratification.

Table 2
The univariable and multivariable Cox regression results of Clinical features.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95 % CI)	*P* value	Hazard ratio (95 % CI)	*P* value
Gensini	61	1.012 (1.002–1.023)	0.022	1.020 (1.004–1.036)	0.012
CREAT	61	1.306 (1.113–1.532)	0.001	1.346 (1.100–1.647)	0.004

the risk of reinfarction and cardiac death [20]. Consequently, *MMP9* holds promise as a valuable diagnostic and prognostic marker for AMI and a potential therapeutic target.

GNLY is an enzyme secreted by natural killer cells and T lymphocytes that can kill target cells and is involved in immune regulation and myocardial protection after AMI [21]. The downregulation of *GNLY* may be due to the impairment or reduction of natural killer cells and cytotoxic lymphocytes function or number caused by STEMI. The low expression of *GNLY* may lead to decreased immune surveillance function, increased infection risk, delayed myocardial repair, etc., affecting the prognosis of STEMI patients. Similarly, *GNLY* may also serve as a valuable diagnostic and prognostic indicator for AMI, with potential therapeutic implications.

Unfortunately, *MMP9* and *GNLY* were not significantly associated with patient prognosis in our study. There may be several explanations for this phenomenon. First, their expression levels may be influenced by other factors, such as age, gender, genetic polymorphism, drug treatment, etc., resulting in reduced association with prognosis. Secondly, the two genes may be only part of the pathogenesis mechanism of STEMI, rather than decisive factors. The prognosis of AMI may be influenced by multiple genes and pathways, such as plasminogen activator inhibitor 1 [22], vascular endothelial growth factor [23], etc. Thirdly, *MMP9* and *GNLY* may only play a role in the early stage of STEMI, rather than continuously affect prognosis. Notably, our study exclusively assessed gene expression levels during the early stages of AMI, without accounting for temporal changes in gene expression over time. Some genes may show association with prognosis only in the late stage of STEMI, such as aging-related genes *ETS2* [24]. In summary, the phenomenon that *MMP9* and *GNLY* are not significantly associated with AMI prognosis may reflect the complexity and dynamics of STEMI pathogenesis and prognostic determinants. It is warranted to further study the mechanism and time effect of them in STEMI, as well as other genes and pathways that may affect prognosis.

Our primary finding centered on the significant association of *S100A12* and *IL2RB* as DEIRGs with STEMI prognosis. S100A12, a member of the S100 protein family, plays a crucial role in regulating inflammatory processes and immune responses [25,26]. Primarily secreted by neutrophils, S100A12 interacts with the receptor for advanced glycation end products, implicated in atherosclerosis and cardiovascular diseases [27,28]. Prior studies consistently reported elevated S100A12 levels in the plasma of STEMI patients, suggesting its potential as an early-stage prognostic biomarker [29]. Our findings align with existing research, demonstrating significant upregulation of S100A12 in the serum of STEMI patients [11,29]. This upregulation may result from neutrophil activation, chemotaxis, and release triggered by STEMI [30]. The binding of S100A12 to RAGE on vascular endothelial cells, monocytes, and macrophages initiates pro-inflammatory responses, including activation of the NF-κB signaling pathway and subsequent expression of various

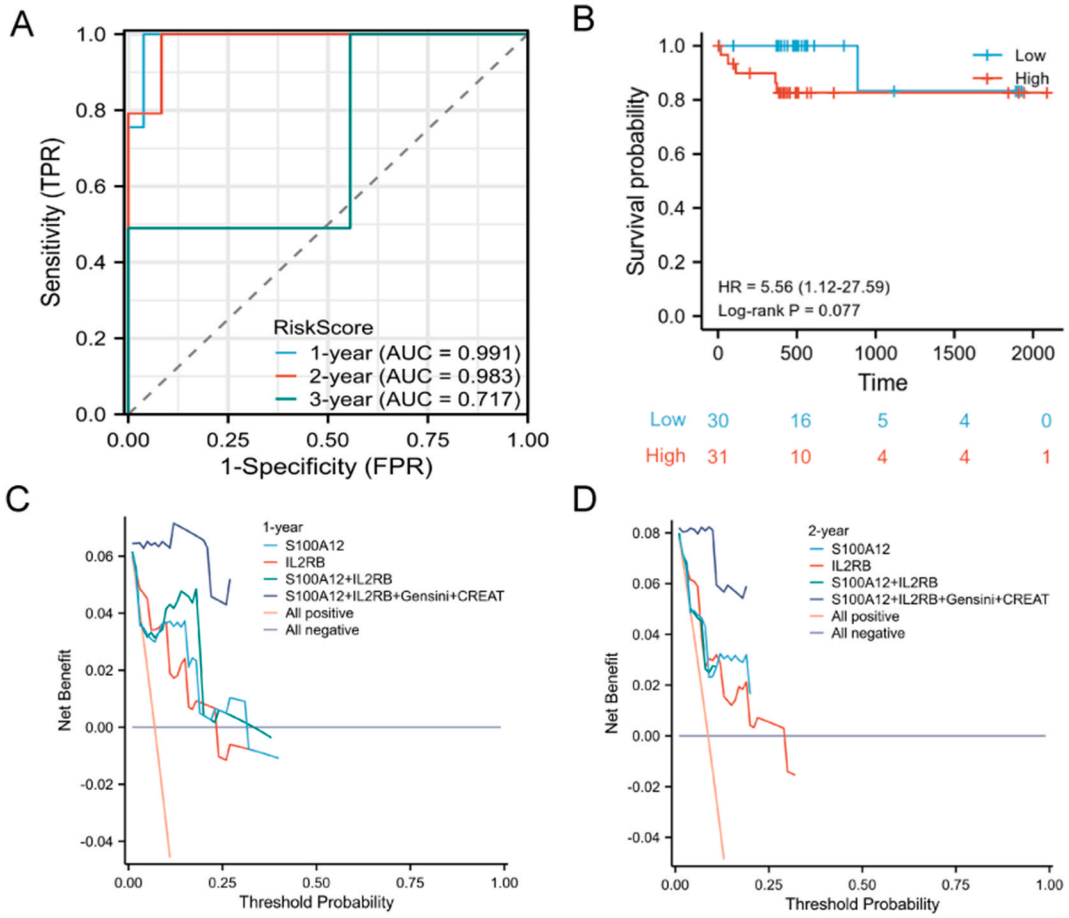

Fig. 5. The signature predicts survival of STEMI patients *in* GSE49925. (A) Time-dependent ROC curves. TPR true positive rate, FPR false positive rate. (B) Kaplan-Meier survival curves. HR hazard ratio. (C&D) Decision curve analysis of the constructed model at one and two years of onset.

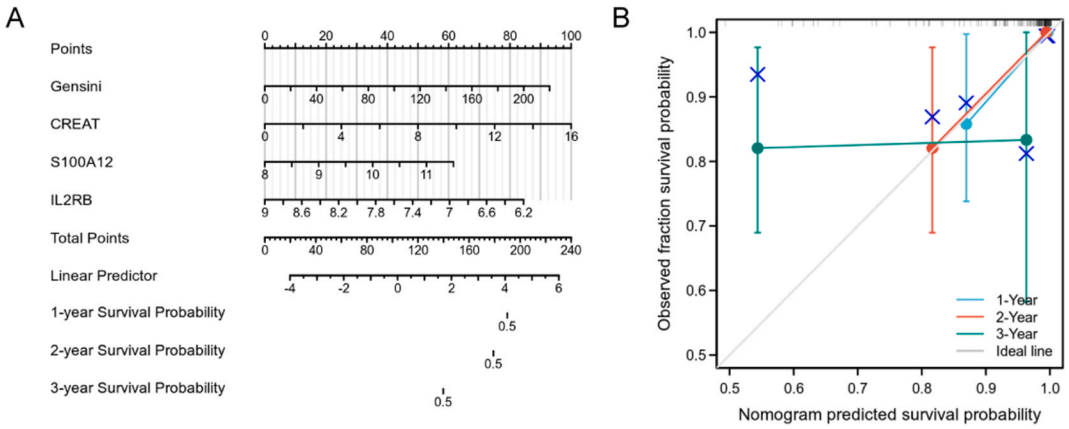

Fig. 6. Construction and performance verification of nomogram. (A) Nomogram for predicting 1-,2-, and 3-year survival in the entire cohort. (B) Calibration curves of nomogram on consistency between predicted and observed 1-, 2- and 3-year survival in the entire cohort. The grey line at 45° implicated a perfect prediction, and the actual performances of our nomogram were shown in colored lines.

cytokines and adhesion molecules [15]. This interaction contributes to the inflammatory milieu characteristic of AMI, promoting further leukocyte recruitment and sustaining the inflammatory response. The high expression of *S100A12* may lead to endothelial cell damage, vascular inflammation, plaque instability, etc., increasing the short-term risk of cardiovascular events in STEMI patients [23].

Table 3
Clinical characteristics of external validation cohort.

Parameter	High risk (n = 46)	Low risk (n = 46)	P value
Age	63.2 ± 10.6	61.5 ± 9.8	0.412
Male [n(%)]	34(73.9 %)	35(76.1)	0.810
Hypertension [n(%)]	28(60.9 %)	26(56.5)	0.672
Diabetes [n(%)]	14(30.4)	12(26.1)	0.643
Heart rate (beats/min)	84.1 ± 13.6	82.3 ± 12.4	0.518
Systolic blood pressure (mmHg)	138.2 ± 19.4	136.5 ± 18.7	0.643
Scr (mg/dL)	1.20(1.08, 1.27)	0.89(0.85, 0.94)	<0.001
Gensini score	42.3 ± 4.6	36.9 ± 8.7	0.016
S100A12($2^{-\Delta\Delta Ct}$)	3.20(2.70,3.63)	1.60(1.30,1.93)	<0.001
IL2RB($2^{-\Delta\Delta Ct}$)	0.00(-0.43,0.50)	2.0(1.70,2.30)	<0.001

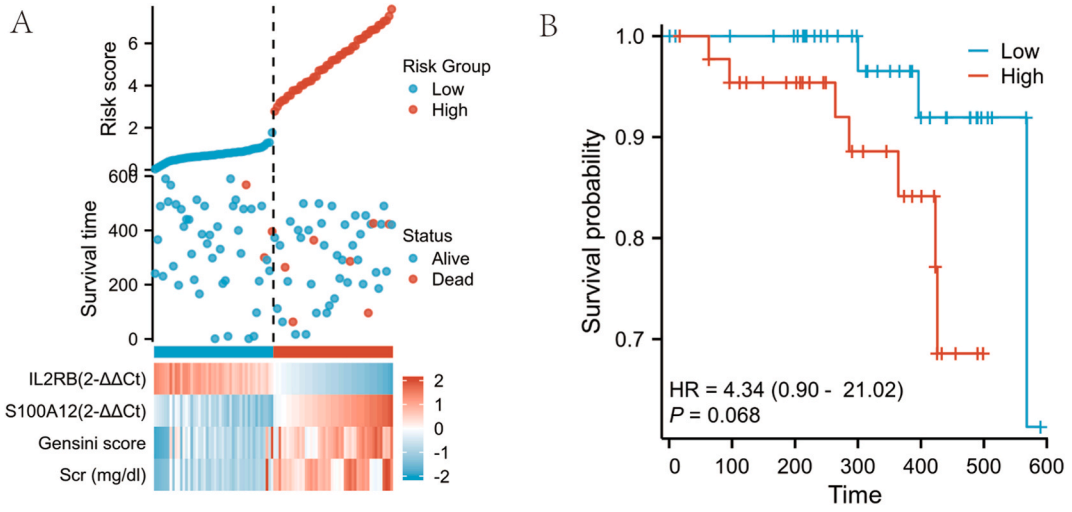

Fig. 7. The results of external validation. (A) A risk distribution plot and (B) Kaplan-Meier survival curves analysis showed significant differences in the incidence of MACEs among different risk groups.

Therefore, *S100A12* may be a valuable prognostic indicator for STEMI, as well as a potential therapeutic target.

IL2RB, identified as the beta subunit of the interleukin-2 receptor (IL-2Rβ), is pivotal in the immune system's reaction to acute tissue damage [31]. Predominantly found on lymphocyte surfaces, IL-2Rβ's activation is instrumental in T-cell growth and diversification [32]. Altering IL2RB expression may impact the mobilization and stimulation of regulatory T cells, which mitigate myocardial damage by curbing rampant inflammation. In AMI scenarios, IL2RB's expression irregularities could signify a shifted immune response, possibly tipping the scales of pro-inflammatory and anti-inflammatory signaling [33]. Our analysis revealed a notable downregulation of IL2RB in STEMI patient serum, aligning with findings from prior studies [11,34]. The downregulation of IL2RB may be due to immune suppression or tolerance caused by STEMI. The low expression of IL2RB may lead to impaired immune function, increased infection risk, delayed myocardial repair, etc., affecting the prognosis [34]. Therefore, IL2RB may also be a valuable prognostic indicator for AMI, as well as a potential therapeutic target.

Further, we established a predictive model composed of S100A12 and IL2RB by logistic regression analysis, combined with serum creatinine level and coronary angiography Gensini score, which can effectively predict the risk of death within two years in STEMI patients. Some peers have developed early diagnostic models for AMI utilizing DEIRGs; however, the differential genes involved differ from those in our study [35,36]. This indicates that the pathogenesis and prognosis of AMI may be associated with a variety of immune-related genes. The expression of these genes can be influenced by various factors during STEMI onset and outcome, and their prognostic associations may also be affected by diverse elements. Notably, Gorav et al. developed biomarker-based prediction models for 1-year risk of cardiovascular death and AMI in acute coronary syndrome patients undergoing PCI [37]. Their model incorporated common clinical variables along with peripheral blood levels of growth differentiation factor 15 and N-terminal pro-B-type natriuretic peptide. Interestingly, there is currently no published literature on constructing prognosis models specifically for STEMI patients using DEIRG. The dysregulation of these genes, in conjunction with clinical indicators of renal function and coronary artery disease severity, offers insights into the systemic nature of AMI and its impact on patient outcomes. The combined analysis underscores the importance of a multi-faceted approach to understanding and predicting the prognosis of AMI, taking into account both molecular and clinical factors.

Importantly, we rigorously validated the accuracy and stability of our constructed predictive model through multidimensional assessments. Firstly, a risk factor distribution plot visually demonstrated significant survival outcome differences across patients with

varying model score intervals, effectively distinguishing different risk levels. Secondly, Kaplan-Meier survival curves revealed substantial survival rate disparities among risk-stratified patients during the 3-year follow-up period, indicating the model's efficacy in short-term survival prediction. Thirdly, time-dependent ROC curves validated the model's predictive value over 1-year, 2-year, and 3-year intervals, with optimal performance observed in the first two years. Decision curve analysis of the clinical application net benefit further confirmed the model's utility in guiding clinical decisions and interventions. Lastly, our prognostic nomogram, calibrated against actual survival rates, effectively assessed individualized patient risk within the initial two years. In summary, these results indicated that our prediction model demonstrates high accuracy, stability, and reliability, providing valuable prognostic information and clinical guidance across different time points and scenarios.

4.1. Clinical utility and application

The practical significance of our research holds considerable importance for primary care physicians, who frequently serve as the initial medical consultants for individuals with cardiovascular conditions. The prognostic framework we've developed, which synthesizes DEIRGs with clinical factors, provides an innovative instrument for the preliminary categorization of STEMI patients based on their risk levels. This tool has the potential to inform clinical decision-making in several ways. Firstly, the signature can aid primary care physicians in identifying patients at higher risk of adverse outcomes post-STEMI. By incorporating this signature into routine evaluations, physicians can prioritize referrals and expedite specialist consultations for those who may benefit from more aggressive treatment or closer monitoring. Secondly, the integration of serum creatinine levels and Gensini scores with DEIRGs provides a more comprehensive view of a patient's condition. This holistic approach not only reflects the cardiac event's severity but also captures the systemic nature of the immune response, which is crucial for tailoring patient management plans. Incorporating this signature into electronic health records could streamline the risk assessment process, allowing for more efficient patient care. Additionally, the prognostic nomogram we constructed can serve as a user-friendly tool for physicians to estimate survival rates, facilitating more informed discussions with patients regarding their prognosis and treatment options.

4.2. Strengths and limitations of the study

This study boasts certain advantages in advancing STEMI patient prognosis through the systematic analysis of public microarray data to pinpoint key DEIRGs. We have developed a predictive model that marries these genes with clinical factors such as serum creatinine and Gensini scores, refined via LASSO regression. Our model has been validated on multiple fronts, including external cohorts, and stands out for its accuracy in forecasting two-year outcomes. This represents a pioneering effort in employing DEIRGs for STEMI prognosis, paving the way for more personalized treatment and nuanced risk assessment.

It's essential to recognize certain constraints within our investigation: (1) The reliance on public databases introduced constraints in data volume, sample source variability, and potential selection bias. (2) The use of microarray technology for gene expression data, while effective, comes with inherent noise and variability that may impact data accuracy and reproducibility. Additionally, this technology is limited to known genes, potentially overlooking significant immune-related genes yet to be annotated. (3) The predictive model was constructed using Cox regression, which carries certain assumptions and parameters that could influence the model's generalizability and interpretability. (4) In the external validation of our predictive model, the Logrank test *P*-value for the Kaplan-Meier survival curves approached statistical significance, indicating a promising trend. Despite the limitation, the multidimensional validation we conducted affirms the model's precision and stability. We are dedicated to conducting further research with larger cohorts and longer follow-up periods to robustly substantiate our findings.

5. Conclusion

This study underscores the pivotal role of IRGs in AMI prognosis. The prognostic signature developed through the integration of DEIRGs and two readily available clinical variables offers a novel approach to stratifying patients based on survival risk. The model was demonstrated to possess high accuracy and stability, demonstrating that the model has higher predictive ability and clinical application value than traditional clinical risk factors or single biomarkers. Additionally, our exploration into the functions of *S100A12* and *IL2RB* in AMI's immune response sheds light on the intricate mechanisms that propel the disease's progression, paving the path for precise therapeutic strategies.

Ethics statement

The ethical review board of Tianjin Chest Hospital granted approval for all human participant-related procedures in this study (approval number 202301636).

Funding

This work was supported by a grant from National Natural Science Foundation of China (82000267), Tianjin Science and Technology Program of China (21JCYBJC01590) and Tianjin Health Research Project (2023013).

Data availability statement

The datasets produced and/or examined in the course of this study can be accessed in the GEO database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE60993, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61144, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE49925).

CRediT authorship contribution statement

Wei Gao: Writing – review & editing, Writing – original draft, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xiaoyan Wang:** Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Xingjie Wang:** Writing – review & editing, Validation, Project administration, Conceptualization. **Lei Huang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We extend our gratitude to Greg Gibson for his generous contribution of additional clinical and follow-up data for the GSE49925 dataset.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e31247.

References

[1] D.M. Tehrani, A.H. Seto, Third universal definition of myocardial infarction: update, caveats, differential diagnoses, Cleve. Clin. J. Med. 80 (12) (2013) 777–786.
[2] D.H. Fitchett, P. Theroux, J.M. Brophy, et al., Assessment and management of acute coronary syndromes (ACS): a Canadian perspective on current guideline-recommended treatment–part 2: ST-segment elevation myocardial infarction, Can. J. Cardiol. 27 (Suppl A) (2011) S402–S412.
[3] A. Koutsoukis, I. Kanakakis, Challenges and unanswered questions in STEMI management, Hellenic J. Cardiol. 60 (4) (2019) 211–215.
[4] Y. Xue, J. Shen, W. Hong, et al., Risk stratification of ST-segment elevation myocardial infarction (STEMI) patients using machine learning based on lipid profiles, Lipids Health Dis. 20 (1) (2021) 48.
[5] L. Zhang, Z. Liu, Y. Zhu, et al., Development and validation of a prognostic model for predicting post-discharge mortality risk in patients with ST-segment elevation myocardial infarction (STEMI) undergoing primary percutaneous coronary intervention (PPCI), J. Cardiothorac. Surg. 19 (1) (2024) 163.
[6] Z. Lin, H. Xu, Y. Chen, X. Zhang, J. Yang, Identification of potential biomarkers and immune-related pathways related to immune infiltration in patients with acute myocardial infarction, Transpl. Immunol. 74 (2022) 101652.
[7] M.A. Zambrano, P. Alcaide, Immune cells in cardiac injury repair and remodeling, Curr. Cardiol. Rep. 25 (5) (2023) 315–323.
[8] X. Zhu, T. Yin, T. Zhang, et al., Identification of immune-related genes in patients with acute myocardial infarction using machine learning methods, J. Inflamm. Res. 15 (2022) 3305–3321.
[9] P.F. Zheng, Q.C. Zou, L.Z. Chen, P. Liu, Z.Y. Liu, H.W. Pan, Identifying patterns of immune related cells and genes in the peripheral blood of acute myocardial infarction patients using a small cohort, J. Transl. Med. 20 (1) (2022) 321.
[10] X. Tan, Q. Dai, H. Sun, et al., Systematic bioinformatics analysis based on public and second-generation sequencing transcriptome data: a study on the diagnostic value and potential mechanisms of immune-related genes in acute myocardial infarction, Front. Cardiovasc. Med. 9 (2022) 863248.
[11] H. Zhai, L. Huang, Y. Gong, et al., Human plasma transcriptome implicates dysregulated S100A12 expression: a strong, early-stage prognostic factor in ST-segment elevated myocardial infarction: bioinformatics analysis and experimental verification, Front. Cardiovasc. Med. 9 (2022) 874436.
[12] J. Kim, N. Ghasemzadeh, D.J. Eapen, et al., Gene expression profiles associated with acute myocardial infarction and risk of cardiovascular death, Genome Med. 6 (5) (2014) 40.
[13] H.J. Park, J.H. Noh, J.W. Eun, et al., Assessment and diagnostic relevance of novel serum biomarkers for early decision of ST-elevation myocardial infarction, Oncotarget 6 (15) (2015) 12970–12983.
[14] Y. Zaidi, E.G. Aguilar, M. Troncoso, D.V. Ilatovskaya, K.Y. DeLeon-Pennell, Immune regulation of cardiac fibrosis post myocardial infarction, Cell. Signal. 77 (2021) 109837.
[15] A. Anzai, S. Ko, K. Fukuda, Immune and inflammatory networks in myocardial infarction: current research and its potential implications for the clinic, Int. J. Mol. Sci. 23 (9) (2022).
[16] J. Goerg, M. Sommerfeld, B. Greiner, et al., Low-dose empagliflozin improves systolic heart function after myocardial infarction in rats: regulation of MMP9, NHE1, and SERCA2a, Int. J. Mol. Sci. 22 (11) (2021).
[17] Y. Tan, X. Ji, Z. Mo, Y. Zhou, Serum YKL-40 positively correlates with MMP-9 and CRP in patients with acute ST segment elevation myocardial infarction following emergency treatment, Medicine 98 (47) (2019) e17950.
[18] A. Villano, G. Barcaro, S. Monti, et al., Molecularly imprinted nanoparticles towards MMP9 for controlling cardiac ECM after myocardial infarction: a predictive experimental-computational chemistry investigation, Biomedicines 10 (9) (2022).
[19] A. Domínguez-Rodríguez, D. Hernández-Vaquero, P. Abreu-González, et al., Early treatment of acute myocardial infarction with melatonin: effects on MMP-9 and adverse cardiac events, J. Clin. Med. 11 (7) (2022).

[20] X. Chen, S. Wang, W. Xu, M. Zhao, Y. Zhang, H. Xiao, Metformin directly binds to MMP-9 to improve plaque stability, J. Cardiovasc. Dev. Dis. 10 (2) (2023).
[21] V. Persic, A. Ruzic, B. Miletic, et al., Granulysin expression in lymphocytes that populate the peripheral blood and the myocardium after an acute coronary event, Scand. J. Immunol. 75 (2) (2012) 231–242.
[22] M. Pavlov, V. Nikolić-Heitzler, Z. Babić, et al., Plasminogen activator inhibitor-1 activity and long-term outcome in patients with ST-elevation myocardial infarction treated with primary percutaneous coronary intervention: a prospective cohort study, Croat. Med. J. 59 (3) (2018) 108–117.
[23] M. Braile, S. Marcella, L. Cristinziano, et al., VEGF-A in cardiomyocytes and heart diseases, Int. J. Mol. Sci. 21 (15) (2020).
[24] J. Xiang, J. Shen, L. Zhang, B. Tang, Identification and validation of senescence-related genes in circulating endothelial cells of patients with acute myocardial infarction, Front. Cardiovasc. Med. 9 (2022) 1057985.
[25] D. Foell, D. Kane, B. Bresnihan, et al., Expression of the pro-inflammatory protein S100A12 (EN-RAGE) in rheumatoid and psoriatic arthritis, Rheumatology 42 (11) (2003) 1383–1389.
[26] J. Goyette, C.L. Geczy, Inflammation-associated S100 proteins: new mechanisms that regulate function, Amino Acids 41 (4) (2011) 821–842.
[27] J. Farokhzadian, P. Mangolian Shahrbabaki, V. Bagheri, S100A12-CD36 axis: a novel player in the pathogenesis of atherosclerosis? Cytokine 122 (2019) 154104.
[28] H. Singh, V. Rai, D.K. Agrawal, LPS and oxLDL-induced S100A12 and RAGE expression in carotid arteries of atherosclerotic Yucatan microswine, Mol. Biol. Rep. 49 (9) (2022) 8663–8672.
[29] Y. Sun, M. Wang, X. Tan, H. Zhang, S. Yang, Identification of oxidative stress-related biomarkers in acute myocardial infarction, Int. J. Gen. Med. 16 (2023) 4805–4818.
[30] Z. Yao, B. Zhang, G. Niu, et al., Neutrophil infiltration characterized by upregulation of S100A8, S100A9, S100A12 and CXCR2 is associated with the Co-occurrence of crohn's disease and peripheral artery disease, Front. Immunol. 13 (2022) 896645.
[31] J. Pugin, How tissue injury alarms the immune system and causes a systemic inflammatory response syndrome, Ann. Intensive Care 2 (1) (2012) 27.
[32] V.A. Peters, J.J. Joesting, G.G. Freund, IL-1 receptor 2 (IL-1R2) and its role in immune regulation, Brain Behav. Immun. 32 (2013) 1–8.
[33] M. Zhang, J. Li, C. Hua, J. Niu, P. Liu, G. Zhong, Exploring an immune cells-related molecule in STEMI by bioinformatics analysis, BMC Med. Genomics 16 (1) (2023) 151.
[34] M. Alderdice, S.G. Craig, M.P. Humphries, et al., Evolutionary genetic algorithm identifies IL2RB as a potential predictive biomarker for immune-checkpoint therapy in colorectal cancer, NAR Genom. Bioinform. 3 (2) (2021) lqab016.
[35] X. Tong, X. Zhao, X. Dang, Y. Kou, J. Kou, Predicting diagnostic gene biomarkers associated with immune checkpoints, N6-methyladenosine, and ferroptosis in patients with acute myocardial infarction, Front. Cardiovasc. Med. 9 (2022) 836067.
[36] J. Liu, L. Chen, X. Zheng, C. Guo, Identification of immune-related genes in acute myocardial infarction based on integrated bioinformatical methods and experimental verification, PeerJ 11 (2023) e15058.
[37] G. Batra, J. Lindbäck, R.C. Becker, et al., Biomarker-based prediction of recurrent ischemic events in patients with acute coronary syndromes, J. Am. Coll. Cardiol. 80 (18) (2022) 1735–1747.