

RESEARCH

Open Access



Diagnostic biomarkers and immune infiltration profiles common to COVID-19, acute myocardial infarction and acute ischaemic stroke using bioinformatics methods and machine learning

Ya-Nan Ma^{1†}, Si-Rong Ma^{2†}, Li Yang¹, Juan Wu¹, Ya-Rong Wang¹, Li-Jia Bao¹, Li Ma¹, Qing-Qiu Wu^{1*} and Zhen-Hai Wang^{3,4,5*}

Abstract

Background COVID-19 is a disease that affects people globally. Beyond affecting the respiratory system, COVID-19 patients are at an elevated risk for both venous and arterial thrombosis. This heightened risk contributes to an increased probability of acute complications, including acute myocardial infarction (AMI) and acute ischemic stroke (AIS). Given the unclear relationship between COVID-19, AMI, and AIS, it is crucial to gain a deeper understanding of their associations and potential molecular mechanisms. This study aims to utilize bioinformatics to analyze gene expression data, identify potential therapeutic targets and biomarkers, and explore the role of immune cells in the disease.

Methods This study employed three Gene Expression Omnibus (GEO) datasets for analysis, which included data on COVID-19, AMI and AIS. We performed enrichment analysis on the co-DEGs for these three diseases to clarify gene pathways and functions, and also examined the relationship between co-DEGs and immune infiltration. Machine learning techniques and protein–protein interaction networks (PPI) were used to identify hub genes within the co-DEGs. Finally, we employed a dual validation strategy integrating independent GEO datasets and in vitro experiments with human blood samples to comprehensively assess the reliability of our experimental findings.

Results We identified 88 co-DEGs associated with COVID-19, AMI and AIS. Enrichment analysis results indicated that co-DEGs were significantly enriched in immune inflammatory responses related to leukocytes and neutrophils. Immune infiltration analysis revealed significant differences in immune cell populations between the disease group and the normal group. Finally, genes selected through machine learning methods included: CLEC4E, S100A12, and IL1R2. Based on the PPI network, the top ten most influential DEGs were identified as MMP9, TLR2, TLR4, ITGAM,

[†]Ya-Nan Ma and Si-Rong Ma contributed equally to this work and share first authorship.

*Correspondence:

Qing-Qiu Wu
wqq9455@163.com
Zhen-Hai Wang
nyfywzh@163.com

Full list of author information is available at the end of the article



S100A12, FCGR1A, CD163, FCER1G, FPR2, and CLEC4D. The integration of the protein–protein interaction (PPI) network with machine learning techniques facilitated the identification of S100A12 as a potential common biomarker for early diagnosis and a therapeutic target for all three diseases. Ultimately, validation of S100A12 showed that it was consistent with our experimental results, confirming its reliability as a biomarker. Moreover, it demonstrated good diagnostic performance for the three diseases.

Conclusion We employed bioinformatics methods and machine learning to investigate common diagnostic biomarkers and immune infiltration characteristics of COVID-19, AMI and AIS. Functional and pathway analyses indicated that the co-DEGs were primarily enriched in immune inflammatory responses related to leukocytes and neutrophils. Through two machine learning approaches and the PPI network, and subsequent validation and evaluation, we identified S100A12 as a potential common therapeutic target and biomarker related to immune response that may influence these three diseases.

Keywords COVID-19, Blood clotting abnormality, Bioinformatics, Biomarker, Machine learning, Immune infiltration

Introduction

The outbreak of the COVID-19 virus in 2019 was a global event that affected nearly all of humanity. Due to the lack of specific symptomatic treatments, it prompted the global medical community to delve deeply into its pathophysiological mechanisms. However, the virus continues to spread globally, driven by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), commonly known as COVID-19. Clinically, it primarily presents as lung infections and various respiratory symptoms. In severe instances, the condition can escalate to acute respiratory distress syndrome (ARDS) [1].

In COVID-19 patients, in addition to respiratory system impairment, researchers have also observed severe complications related to the cardiovascular and cerebrovascular systems [2, 3]. COVID-19 patients face an elevated risk of venous and arterial thrombosis, which in turn raises the likelihood of acute complications, including acute myocardial infarction (AMI) and acute ischemic stroke (AIS) [4, 5]. Many studies have indicated an increased incidence of AMI and AIS in the COVID-19 population, particularly among the elderly, who exhibit elevated levels of D-dimer and abnormal fibrinogen levels. This group also has a higher mortality rate compared to other populations [2, 6]. The damage to the cardiovascular and cerebrovascular systems is associated with coagulation abnormalities leading to thrombosis. The mechanisms may involve inflammation-induced endothelial injury, the release of cytokines, activation of tissue factor (TF), complement system activation, and the presence of antiphospholipid antibodies [7]. This coagulation abnormality in the COVID-19 population, particularly among the elderly, is prone to result in the occurrence of acute diseases, especially those associated with poor prognosis [2, 3, 8].

Given the ambiguous connections between COVID-19, acute myocardial infarction (AMI), and acute ischemic stroke (AIS), it is essential to explore their relationships

and the underlying molecular mechanisms more thoroughly. Although numerous studies have explored the relationship between COVID-19 and coagulation-related disorders, the existing literature is largely confined to analyses of single diseases or associations between two diseases, lacking a comprehensive exploration of the common mechanisms underlying all three conditions. This study aims to utilize bioinformatics to analyze gene expression data, exploring the common gene expression changes among these diseases and identifying potential therapeutic targets and biomarkers. Through this research, we hope to provide scientific evidence for the treatment and prevention strategies of COVID-19 and its related complications, thereby improving clinical outcomes for patients and reducing disease risk.

This study integrated three Gene Expression Omnibus (GEO) datasets for analysis: GSE171110, GSE66360, and GSE58294. Initially, we employed bioinformatics methods to identify differentially expressed genes (DEGs) in each dataset and to pinpoint the common DEGs (co-DEGs) among these three diseases. Subsequently, we conducted enrichment analysis on the co-DEGs to clarify gene pathways and functions, and examined the relationship between co-DEGs and immune infiltration. Finally, machine learning techniques and protein–protein interaction networks (PPI) were used to identify hub genes within the co-DEGs. Our research provides new perspectives for the early diagnosis and treatment of the three diseases. Compared to existing studies, this work not only fills the gap in research on common mechanisms across multiple diseases but also lays a theoretical foundation for developing combined therapeutic strategies targeting COVID-19, AMI, and AIS.

Materials and methods

Collection of the datasets

To analyze the potential relationships and therapeutic targets among COVID-19, AMI, and AIS, we utilized

three GEO databases to obtain microarray and RNA-seq datasets. The GEO dataset for COVID-19 is GSE171110 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE171110>), which includes transcriptomic profiles from 44 COVID-19 samples and 10 healthy control samples. GSE171110 is based on the Illumina HiSeq 2500 (Homo sapiens) (GPL16791) platform for RNA sequencing analysis. The GEO dataset for AMI is GSE66360 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66360>), comprising trans-cryptomic profiles from 49 AMI patients and 50 healthy control samples, based on the Affymetrix Human Genome U133 Plus 2.0 Array (GPL570) platform. Similarly, the dataset for AIS is GSE58294 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58294>), which includes gene expression profiles from 69 patients with cardioembolic stroke (CES) and 23 healthy individuals. Table 1 provides the essential information for the three datasets. We also selected three additional datasets as validation datasets. The COVID-19 validation dataset is GSE152418, which includes transcriptomic profiles of 17 COVID-19 samples and 17 healthy control samples. The AMI validation dataset is GSE48060, which includes transcriptomic profiles of 31 AMI samples and 21 healthy control samples. Similarly, the AIS dataset is GSE16561, which includes gene expression profiles of 39 AIS patients and 24 healthy controls.

Identification of the DEGs and co-DEGs for COVID-19, AMI and AIS.

Based on the research design, the samples were divided into three groups: COVID-19, AMI, and AIS. The R package limma was used to perform differential expression analysis for each group, with a threshold for differentially expressed genes (DEGs) set at $\log_2(\text{Fold Change}) > \pm 1.5$ and $p \text{ value} < 0.05$. DEGs were classified into upregulated and downregulated genes according to the criteria of $\log_2(\text{Fold Change}) > 1.5$ and $\log_2(\text{Fold Change}) < -1.5$, respectively [9].

To identify co-DEGs among the three groups, we intersected the upregulated and downregulated DEGs from each group. A Venn diagram was then generated to visualize the overlapping DEGs. The R package ggplot2 was used to create volcano plots, while the VennDiagram

package was utilized to generate the Venn diagram [10, 11].

GO and KEGG enrichment analysis

To determine the potential functions and pathways of the co-DEGs, we performed GO and KEGG analyses and predictions [12, 13]. GO can predict the possible functions of genes, and after screening the co-DEGs, we classified them according to three criteria: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC). The KEGG database includes pathways that represent molecular interactions, reactions, and relationships, allowing enrichment analysis to explain the functional profiles of co-DEGs from the perspective of influencing specific biological pathways. We employed the R package clusterProfiler for GO and KEGG pathway enrichment of the co-DEGs [14].

Immune infiltration analysis

To assess the composition and abundance of 22 immune cell types in the gene expression data, we utilized the CIBERSORT method, which employs linear support vector regression to deconvolute the transcriptomic expression matrix [15]. This approach explored whether the abundance of relevant immune cells increased or decreased during the progression of the diseases compared to normal individuals, aiding in further investigations into the relationships among COVID-19, AMI, and AIS.

Identification of hub genes

To investigate early biomarkers and potential therapeutic targets for COVID-19, AMI, and AIS, we conducted random forest and XGBoost analyses [16–18]. Random forest is an ensemble algorithm among various machine learning methods, based on multiple decision trees [19]. It predicts the final outcome by aggregating the results of each decision tree, resulting in high accuracy and reliability. In this study, we implemented the Random Forest algorithm using the rfPermute package. We set the number of decision trees (ntree) to 1000 to enhance the model's stability and calculated feature importance based on accuracy decrease and Gini impurity by setting $\text{importance} = \text{TRUE}$. To evaluate the statistical significance

Table 1 The three datasets and their characteristics and analysed measurements in this study

Disease name	GEO accession	GEO platform	Total DEGs count	Up regulated DEGs count	Down regulated DEGs count
COVID-19	GSE171110	GPL16791	6588	3090	3498
AMI	GSE66360	GPL570	1632	833	799
CES	GSE58294	GPL570	824	754	70

of feature importance, we performed 1000 permutation tests ($nrep = 1000$). To address sample imbalance, we employed stratified sampling during the training process, ensuring proportional representation of each class in the training set. On the other hand, XGBoost is a predictive algorithm that integrates small predictive models generated at each step into a larger predictive model. For the XGBoost algorithm, we utilized the XGBoost package in R with key parameters tailored for multi-class classification. We chose a tree-based model (`booster = 'gbtree'`) and limited the maximum depth of each tree to 6 (`max_depth = 6`) to prevent overfitting. The learning rate (`eta`) was set to 0.5 to balance convergence speed and accuracy. Additionally, to further identify hub genes, we examined the known and predicted interactions between proteins in the STRING database (version 12.0, <https://cn.string-db.org>) and used this information to construct a PPI network involving the co-DEGs. After exploring the protein interaction relationships, we visualized the network using Cytoscape software (version 3.10.3, <https://cytoscape.org>) and selected the top 10 nodes as hub genes [20].

Verification and Evaluation of Hub Gene

After preprocessing the data from the three validation datasets, we used the `ggboxplot` and `ggpubr` packages in R to create box plots, visually illustrating the expression differences of S100 A12 between the disease groups and the control groups [21, 22]. To evaluate the diagnostic performance of S100 A12 as a biomarker for the three diseases, we performed ROC curve analysis using the `pROC` package in R and calculated the area under the curve (AUC) along with its 95% confidence interval (CI) [23].

Human Blood Sample Collection

Blood samples from 5 patients each with COVID-19, AMI, and AIS hospitalized at the Cardiovascular and Cerebrovascular Hospital of the General Hospital of Ningxia Medical University in March 2025 were collected. The samples were stored at -80°C prior to extraction. This study was approved by the Medical Research Ethics Review Committee of the General Hospital of Ningxia Medical University (Approval No.: KYLL-2025-0941). The basic information of the patients is provided in Table 2.

QRT-PCR

Total RNA extraction was performed using TRIzol reagent, followed by quantification and purity assessment of the extracted RNA. Subsequently, RNA was reverse-transcribed into cDNA according to the manufacturer's protocol. Using cDNA as the template, QRT-PCR experiments were conducted to monitor the expression

Table 2 Basic information on patients used for QRT-PCR validation

Disease name	Gender	Age	Complications	Time of onset(day)
AMI-1	Male	58	CHD、Hypertension	5
AMI-2	Male	49	CHD	3
AMI-3	Female	38	CHD	4
AMI-4	Male	58	Hypertension, T2DM	1
AMI-5	Female	60	CHD, Hypertension	1
AMI-6	Male	72	CHD	1
COVID-19-1	Male	80	CHD, Hypertension	5
COVID-19-2	Male	79	Hypertension	1
COVID-19-3	Female	70	Pneumonia, CHD	4
COVID-19-4	Male	69	Hypertension, T2DM	8
COVID-19-5	Male	91	Pneumonia, CHD	5
AIS-1	Male	57	Non	2
AIS-2	Male	80	T2DM	3
AIS-3	Female	60	Hypertension, T2DM	8
AIS-4	Male	82	Hypertension	5
AIS-5	Male	62	T2DM	1

CHD: Coronary Heart Disease; T2DM: Type 2 Diabetes Mellitus

Table 3 Primer sequences for QRT-PCR

Primer Name	Primer Sequence (5'to3')
S100 A12-F	GAGCATCTGGAGGGAATTGT
S100 A12-R	CTGCTTCAGCTCACCTTAG
18S rRNA-F	GTAACCCGTTGAACCCATT
18S rRNA-R	CCATCCAATCGGTAGTAGCG

levels of target genes through real-time fluorescence signals. Data analysis was performed using the $2^{-\Delta\Delta\text{CT}}$ method. The primer sequences used in the experiments are detailed in Table 3.

Statistical analysis

In this study, we conducted statistical analyses using R. Independent sample t-tests were used to determine statistical significance between groups. We performed all bioinformatics analyses with R packages of R software. Statistics were considered significant when $P < 0.05$, whereas $P > 0.05$ was not considered significant [24].

Result

Identification of differentially expressed genes

According to our criteria for screening DEGs (\log_2 Fold Change > 1.5 and adjusted p value < 0.05), the COVID-19 dataset GSE171110 identified 6,588 DEGs, including 3,090 upregulated genes and 3,498 downregulated genes. In GSE66360, 1,632 DEGs were identified (833

upregulated and 799 downregulated), while GSE58294 identified 824 DEGs (754 upregulated and 70 downregulated). Figure 1 shows volcano plots for COVID-19, AMI, and AIS, with red and blue points representing upregulated and downregulated genes, respectively (Fig. 1A–C). We identified 88 common DEGs among GSE171110, GSE66360, and GSE58294 (Fig. 1D). The results of the differential expression analysis suggest that there are certain shared mechanisms and interconnections among COVID-19, AMI, and AIS.

GO and KEGG enrichment analysis

GO analysis includes Biological Process (BP), Cellular Component (CC) and Molecular Function (MF), with the GO database selected as the annotation source. Figure 2A displays the top 5 GO terms for BP, CC and MF. Table 4 presents the top 10 terms in each category for BP, CC and MF. The DEGs were significantly enriched in immune inflammatory responses related to leukocytes (BP), granules and granule membranes (CC),

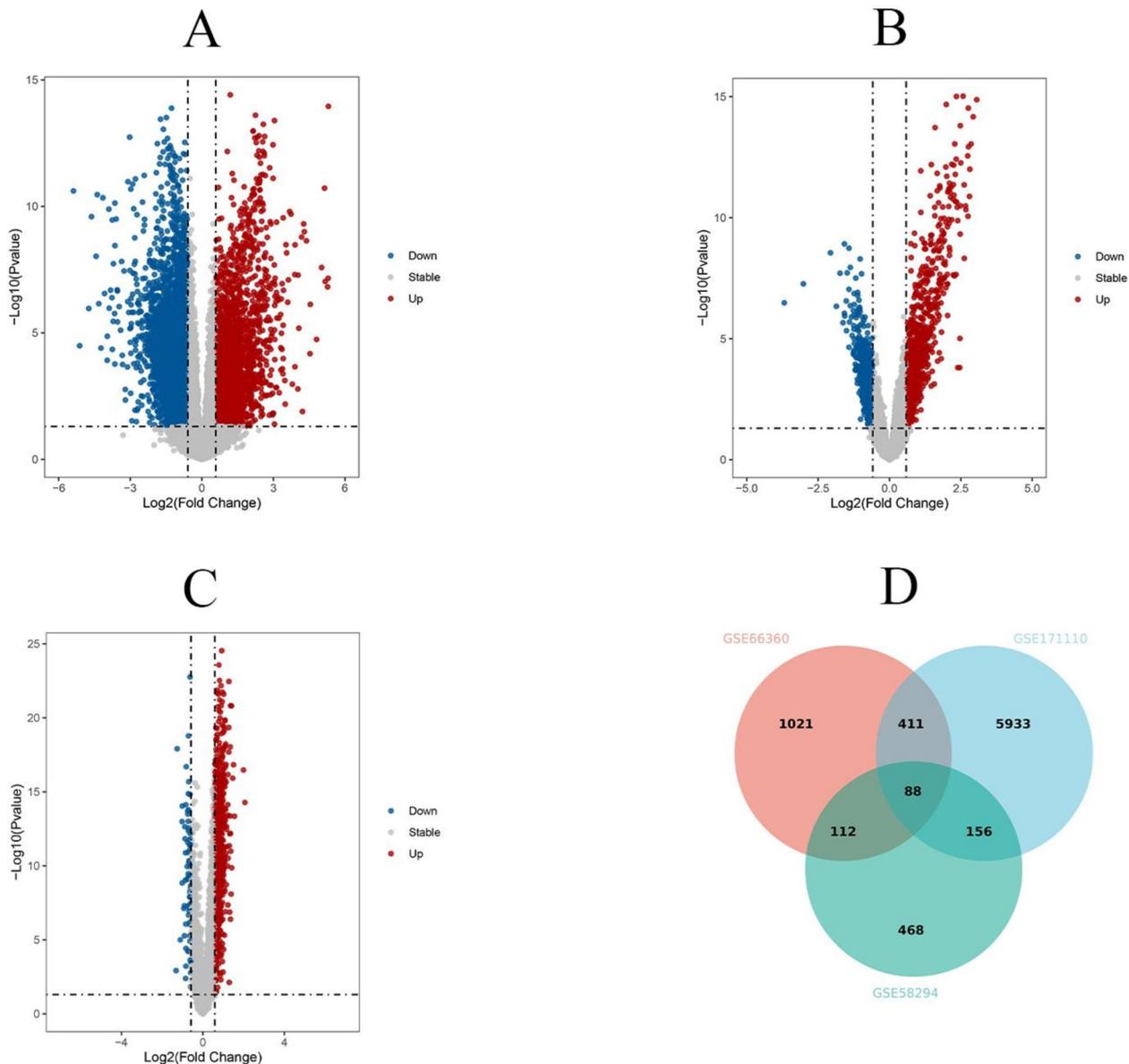


Fig. 1 Volcano plots exhibit DEGs of (A)COVID-19, (B)AMI and (C)AIS (CES). Red for up-regulated and green for down-regulated. (D)The Venn diagram depicts the co-DEGs among GSE171110(COVID-19), GSE66360(AMI) and GSE58294(AIS). The results showed that they had a total of 88 co-DEGs

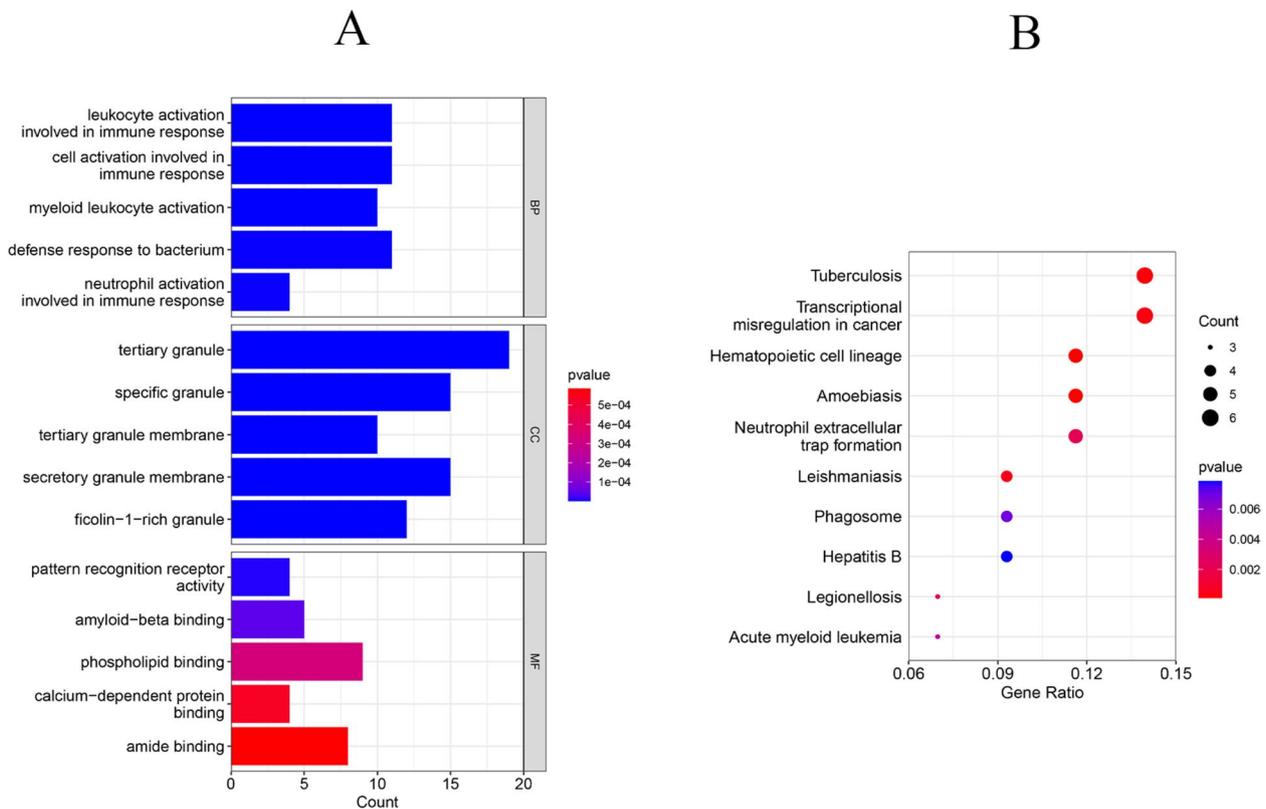


Fig.2 (A)The bar graphs of the ontological analysis of the co-DEGs among COVID-19, AMI and AIS(CES). BP, biological progress; CC, cellular component; MF, molecular function. (B)Bubble graphs indicate the results for KEGG analysis based on the co-DEGs among COVID-19, AMI and AIS

and receptors (MF). These enrichments are linked to immune therapy-related functions.

KEGG pathway analysis identified the top 10 pathways, including tuberculosis, transcriptional dysregulation in cancer, hematopoietic cell lineage, amoebiasis, formation of neutrophil extracellular traps, leishmaniasis, phagosome, hepatitis B, legionellosis, and acute myeloid leukemia. The bubble chart in Fig. 2B lists the potentially enriched pathways for the co-DEGs. For a more detailed illustration, the pathway enrichment analysis is presented in Table 5.

Immune infiltration

Applying the genetic data from the three groups of COVID-19, AMI and AIS, the correlation between 22 types of immune cells and the disease group and the normal group was calculated by the CIBERSORT algorithm, respectively. Figure 3A shows a box-and-line plot comparing immune cell expression between the COVID-19 (yellow) and HC (blue) groups. Certain immune cells, such as Dendritic.cells.resting, Neutrophils, T.cells.CD4.memory.resting, Macrophages.M0, and B.cells.memory, showed significantly different infiltration proportions in the COVID-19 group compared to the HC group.

These cells may play key roles in COVID-19 pathology. Figure 3B is a stacked bar graph to show the composition of immune cells in HC group and disease group. Different colours represent different types of immune cells. Observing the significant difference in the composition of immune cells in HC and COVID-19 groups, certain immune cells may play a more prominent role in the COVID-19 immune response. Figure 4 presents an immune heatmap matrix illustrating the correlations between various types of immune cells. In this matrix, red signifies a positive correlation, with darker shades indicating stronger correlations. Conversely, blue represents negative correlations, where darker hues indicate a more substantial negative relationship. White indicates a weak or no significant correlation. From the graph, it can be concluded that there are significant positive correlations between certain immune cells, such as Mast cells activated and Neutrophils, T cells follicular helper and NK cells activated. There is a significant negative correlation between certain immune cells, for example, Dendritic cells resting and Neutrophils, Plasma cells and T cells CD4 memory resting show a significant negative correlation. The synergistic and antagonistic roles of these cells in the immune response together participate in disease-specific

Table 4 Ontological analysis of co-DEGs among COVID-19, AMI and AIS (CES)

GO ID	ONTOLOGY	Description	p.adjust	Genes	Ggg genelD
GO:0002366	BP	leukocyte activation involved in immune response	5.53E-05		GAB2/CLEC4D/FCER1G/DYSF/CLEC4E/ANXA3/CD177/ITGAM/SEMA4 A/TLR4/CCR6
GO:0002263	BP	cell activation involved in immune response	5.53E-05		GAB2/CLEC4D/FCER1G/DYSF/CLEC4E/ANXA3/CD177/ITGAM/SEMA4 A/TLR4/CCR6
GO:0002274	BP	myeloid leukocyte activation	5.53E-05		S100 A12/TLR2/GAB2/CLEC4D/FCER1G/DYSF/ANXA3/CD177/ITGAM/TLR4
GO:0042742	BP	defense response to bacterium	0.000417022		S100 A12/TLR2/CLEC4D/FCER1G/CLEC4E/ANXA3/FPR2/FCGR1 A/SLPI/TLR4/HP
GO:0002283	BP	neutrophil activation involved in immune response	0.000487054		FCER1G/ANXA3/CD177/ITGAM
GO:0002888	BP	positive regulation of myeloid leukocyte mediated immunity	0.000505654		FCGR1 A/CD177/ARG1/ITGAM
GO:0002699	BP	positive regulation of immune effector process	0.00058543		GAB2/SLC7 A5/CD55/FCGR1 A/FFAR2/CD177/ARG1/ITGAM/TLR4
GO:0002275	BP	myeloid cell activation involved in immune response	0.000964429		GAB2/FCER1G/DYSF/ANXA3/CD177/ITGAM
GO:0002697	BP	regulation of immune effector process	0.001248229		IRAK3/GAB2/SLC7 A5/CD55/FCGR1 A/FFAR2/CD177/ARG1/ITGAM/TLR4
GO:0002886	BP	regulation of myeloid leukocyte mediated immunity	0.001248229		GAB2/FCGR1 A/CD177/ARG1/ITGAM
GO:0070820	CC	tertiary granule	3.03E-20		QPCT/CLEC4D/FCER1G/LRG1/MMP9/MGAM/CD55/MCEMP1/FPR2/SLC2 A3/SIGLECS5/TNFAIP6/TIMP2/ORM1/CD177/ITGAM/CYSTM1/TCN1/HP
GO:0042581	CC	specific granule	1.48E-14		QPCT/CLEC4D/LRG1/ANXA3/MCEMP1/FPR2/SLC2 A3/TIMP2/ORM1/CD177/ARG1/SLPI/ITGAM/TCN1/HP
GO:0070821	CC	tertiary granule membrane	2.84E-11		CLEC4D/FCER1G/MGAM/MCEMP1/FPR2/SLC2 A3/SIGLECS5/CD177/ITGAM/CYSTM1
GO:0030667	CC	secretory granule membrane	1.47E-10		TLR2/CLEC4D/FCER1G/MGAM/CD55/VNN1/MCEMP1/FPR2/SLC2 A3/MME/SIGLECS5/APLP2/CD177/ITGAM/CYSTM1
GO:0101002	CC	ficolin-1-rich granule	6.06E-10		QPCT/CLEC4D/FCER1G/LRG1/MMP9/MGAM/CD55/FPR2/SLC2 A3/SIGLECS5/TNFAIP6/TIMP2
GO:1,904,724	CC	tertiary granule lumen	2.03E-09		QPCT/LRG1/MMP9/TNFAIP6/TIMP2/ORM1/TCN1/HP
GO:0035580	CC	specific granule lumen	4.72E-09		QPCT/LRG1/TIMP2/ORM1/ARG1/SLPI/TCN1/HP
GO:0101003	CC	ficolin-1-rich granule membrane	1.37E-07		CLEC4D/FCER1G/MGAM/CD55/FPR2/SLC2 A3/SIGLECS5
GO:0034774	CC	secretory granule lumen	1.86E-07		S100 A12/S100P/QPCT/LRG1/SRGN/RNASE2/TIMP2/ORM1/ARG1/SLPI/TCN1/HP
GO:0060205	CC	cytoplasmic vesicle lumen	1.86E-07		S100 A12/S100P/QPCT/LRG1/SRGN/RNASE2/TIMP2/ORM1/ARG1/SLPI/TCN1/HP
GO:0038187	MF	pattern recognition receptor activity	0.001765518		TLR2/CLEC4D/CLEC4E/TLR4
GO:0001540	MF	amyloid-beta binding	0.005888223		TLR2/FPR2/ADRB2/ITGAM/TLR4
GO:0005543	MF	phospholipid binding	0.026361914		GAB2/DYSF/ANXA3/WDFY3/KCNJ2/MME/RPH3 A/ARHGAP26/SBF2
GO:0048306	MF	calcium-dependent protein binding	0.026361914		S100 A12/S100P/ANXA3/CD177
GO:0033218	MF	amide binding	0.026361914		TLR2/SLC7 A5/FPR2/MME/FKBP5/ADRB2/ITGAM/TLR4
GO:0140375	MF	immune receptor activity	0.026361914		IL1R2/FCER1G/FPR2/FCGR1 A/CCR6
GO:0042277	MF	peptide binding	0.030593111		TLR2/SLC7 A5/FPR2/MME/ADRB2/ITGAM/TLR4
GO:0019864	MF	IgG binding	0.036426074		FCER1G/FCGR1 A
GO:0055056	MF	D-glucose transmembrane transporter activity	0.038738178		SLC2 A14/SLC2 A3
GO:0030246	MF	carbohydrate binding	0.044172147		CLEC4D/MGAM/CLEC4E/ASGR2/SLC2 A3/SIGLECS5

Table 5 Pathway enrichment analysis of co-DEGs among COVID-19, AMI and AIS (CES)

ID	Description	p.adjust	genelD
hsa04640	Hematopoietic cell lineage	0.010279082	7850/1604/4311/2209/3684
hsa05146	Amoebiasis	0.010279082	7850/7097/383/3684/7099
hsa05152	Tuberculosis	0.011102111	7097/2207/26253/2209/3684/7099
hsa05202	Transcriptional misregulation in cancer	0.01214225	7850/1647/597/4318/2209/3684
hsa05140	Leishmaniasis	0.015823732	7097/2209/3684/7099
hsa04613	Neutrophil extracellular trap formation	0.055315762	7097/2358/2209/3684/7099
hsa05134	Legionellosis	0.055315762	7097/3684/7099
hsa05221	Acute myeloid leukemia	0.084037187	597/2209/3684
hsa04145	Phagosome	0.119488133	7097/2209/3684/7099
hsa05161	Hepatitis B	0.122523639	7097/9586/4318/7099
hsa05215	Prostate cancer	0.158636996	7850/9586/4318
hsa05150	Staphylococcus aureus infection	0.158636996	2358/2209/3684
hsa04064	NF-kappa B signaling pathway	0.158636996	1647/597/7099
hsa04625	C-type lectin receptor signaling pathway	0.158636996	338,339/2207/26253
hsa05144	Malaria	0.25395003	7097/7099
hsa04915	Estrogen signaling pathway	0.289515945	9586/4318/2289
hsa00561	Glycerolipid metabolism	0.34307639	84,649/2713
hsa05321	Inflammatory bowel disease	0.34307639	7097/7099
hsa04664	Fc epsilon RI signaling pathway	0.34307639	9846/2207
hsa04924	Renin secretion	0.34307639	3759/154

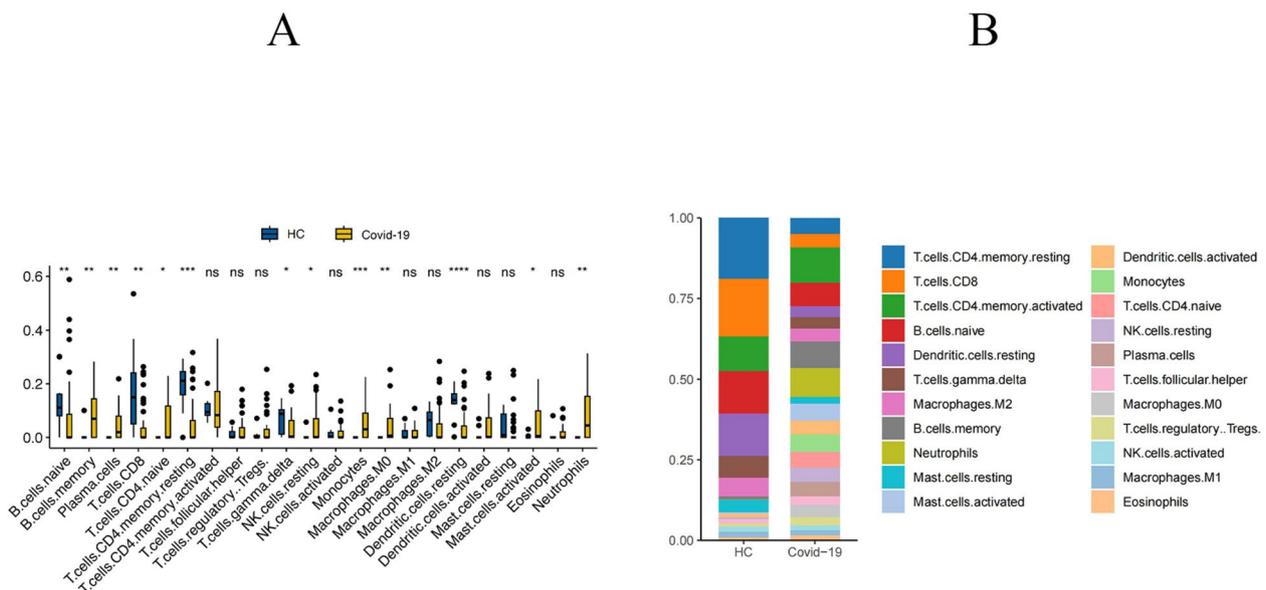


Fig.3 (A) Boxplots were used to show the proportional distribution of different types of immune cells in HC group and COVID-19 group. The results showed that 13/22 immune cells were significantly different between the COVID-19 and HC groups. (B) Stacked bar graphs were used to demonstrate the immune cell composition of HC group and COVID-19 group. The results showed a large difference in immune cell levels between the two groups. *P* values were showed as: *, *p* < 0.05; **, *p* < 0.01; ***, *p* < 0.001; ****, *p* < 0.0001

immunoregulatory processes. Similarly, Fig. 5A shows significantly different and highly statistically significant proportions of T.cells.CD4.memory.resting, Neutrophils and T.cells.gamma.delta as compared to the HC group.

Fig. 5B shows a significantly different composition of immune cells in the HC and AMI groups. Figure 6 shows a significant positive correlation between B cells naive and T cells CD8, Macrophages M2 and Neutrophils.

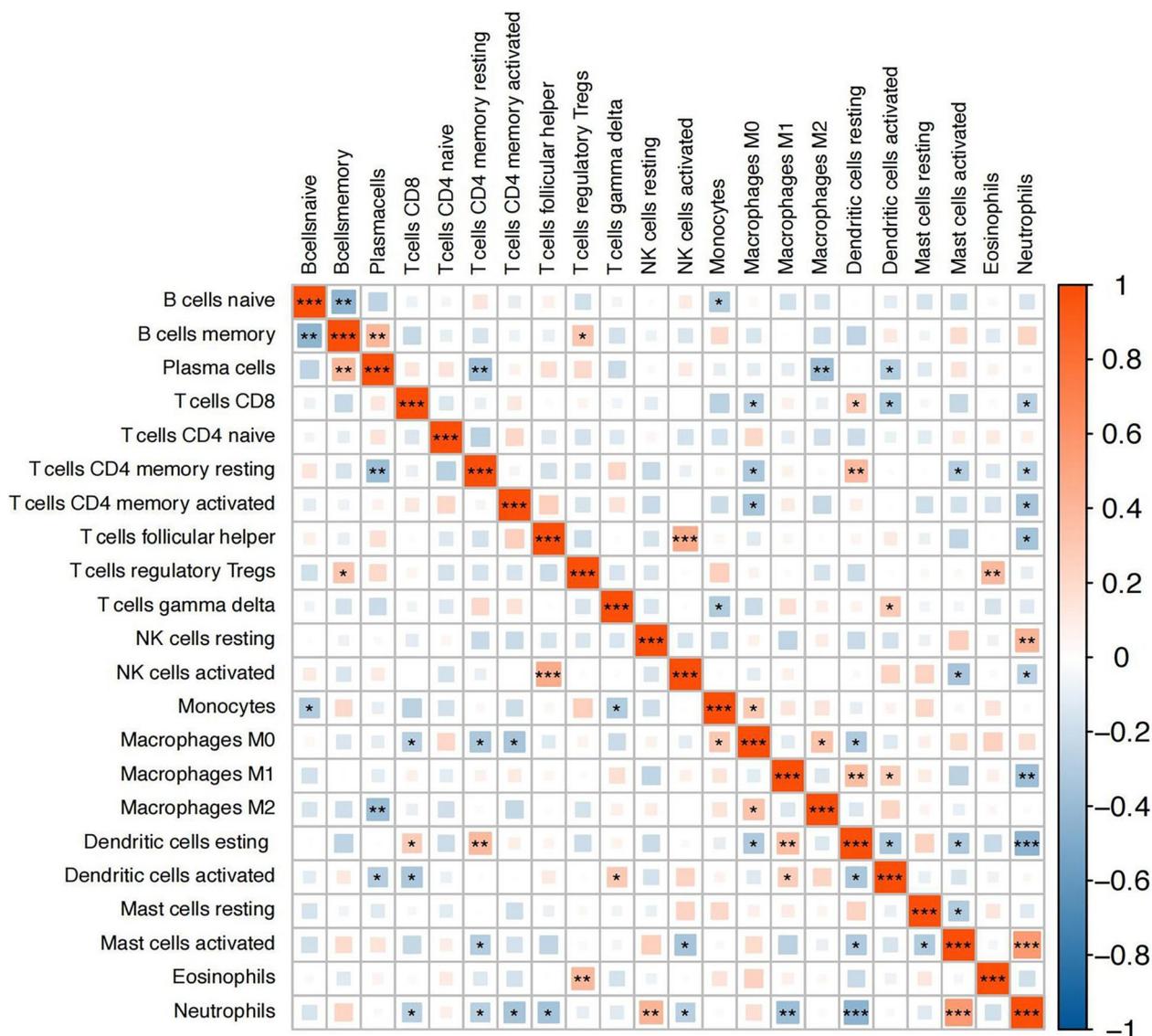


Fig.4 The heatmap matrix is used to demonstrate the correlation between different types of immune cells. The colors and significance markers reflect the strength and significance of the correlation between the cells. Red: positive correlation, the darker the color, the stronger the correlation. Blue: negative correlation, the darker the color, the stronger the negative correlation. White: weak or no significant correlation. *P* values were showed as: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$

There is a significant negative correlation between Neutrophils and T cells gamma delta. Figure 7A shows significantly different proportions of B.cells.naive, T.cells.CD8, and Neutrophils compared to the HC group. Figure 7B shows significantly different immune cell compositions in HC and AIS groups. Immune cell composition was significantly different. Fig. 8 showed a significant positive correlation between Macrophages M2 and Plasma cells, T cells CD8 and T cells regulatory Tregs. There was a significant negative correlation between T cells CD8 and Neutrophils.

Identification of hub genes

Figure 9 illustrates the PPI network analysis results, highlighting the top 10 most influential DEGs. These genes are MMP9, TLR2, TLR4, ITGAM, S100 A12, FCGR1 A, CD163, FCER1G, FPR2, and CLEC4D. We employed two machine learning methods to further identify common early diagnostic biomarkers and potential therapeutic targets for COVID-19, AMI, and AIS. These methods included random forest and XGBoost analyses (Figs .10, 11, 12).

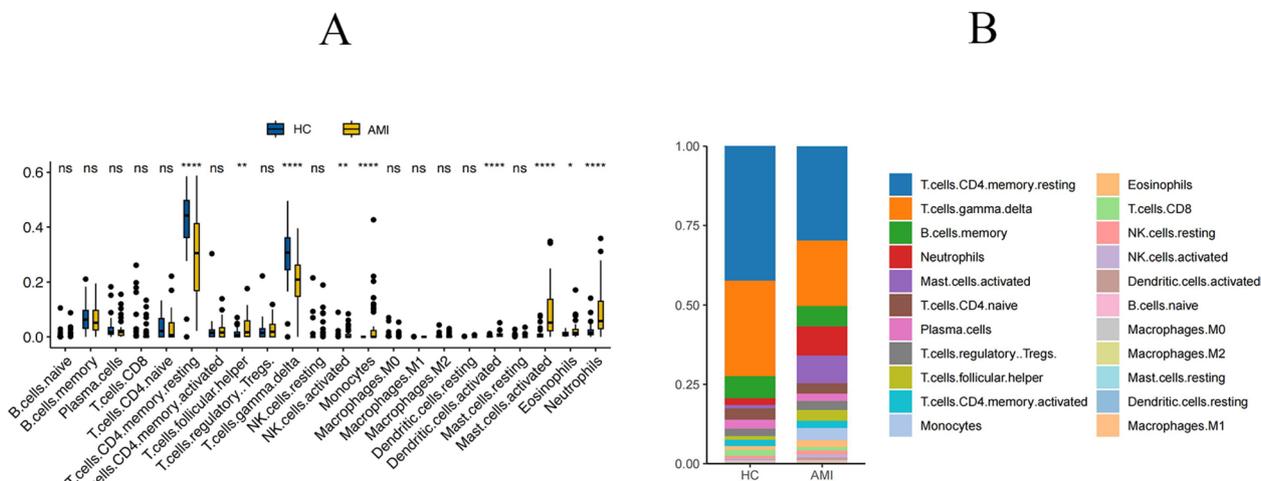


Fig.5 (A) Boxplots were used to show the proportional distribution of different types of immune cells in HC group and AMI group. The results showed that 9/22 immune cells were significantly different between the AMI and HC groups. (B) Stacked bar graphs were used to demonstrate the immune cell composition of HC group and AMI group. The results showed a large difference in immune cell levels between the two groups. *P* values were showed as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

Figure 13 illustrates that in the GSE171110 dataset, random forest selected 74 genes while XGBoost identified 13 genes. In the GSE66360 dataset (Fig. 11), random forest selected 82 genes, with XGBoost identifying 37 genes. For the GSE58294 dataset (Fig. 12), random forest selected 80 genes and XGBoost identified 12 genes. By employing a cross-method approach, we derived the co-DEGs: CLEC4E, S100 A12, and IL1R2.

These hub genes have the potential to act as biomarkers and targets for novel therapeutic strategies in disease research. Ultimately, integrating the PPI network and machine learning methods, we identified S100 A12 as a likely common early diagnostic biomarker and potential therapeutic target for these three diseases.

Verification and Evaluation of Hub Gene

As shown in Fig. 14A, the expression of S100 A12 in the COVID-19 patient group was significantly higher than in the healthy control group, with a statistically significant difference ($P < 0.05$). As shown in Fig. 14B, the expression of S100 A12 in the AMI patient group was significantly higher than in the healthy control group, with a highly statistically significant difference ($P < 0.001$). As shown in Fig. 14C, the expression of S100 A12 in the AIS patient group was significantly higher than in the healthy control group, with a highly statistically significant difference compared to the healthy control group ($P < 0.0001$).

The results of evaluating the diagnostic performance of S100 A12 are as follows: As shown in Fig. 15A, the area under the curve (AUC) for S100 A12 in diagnosing COVID-19 was 0.775(95% CI: 0.608–0.942), indicating moderate diagnostic ability. As shown in Fig. 15B, the

AUC for S100 A12 in diagnosing AMI was 0.782(95% CI: 0.653–0.911), indicating good diagnostic ability. As shown in Fig. 15C, the AUC for S100 A12 in diagnosing AIS was 0.812(95% CI: 0.702–0.922), indicating strong diagnostic ability.

In conclusion, S100 A12 was significantly up-regulated in patients with COVID-19, AMI and AIS, consistent with the experimental results. ROC curve analyses showed that S100 A12 had moderate to good diagnostic efficacy in discriminating between diseased patients and healthy controls, with the best performance in AIS (AUC = 0.812).

QRT-PCR

To further validate the expression characteristics of S100 A12 in different disease states, we used QRT-PCR to measure the expression levels of S100 A12 in patients with COVID-19, AMI, and AIS, as well as in healthy individuals. As shown in Fig. 16, compared to the healthy control group, S100 A12 exhibited an up-regulated expression pattern in all three disease groups. The results are consistent with experimental findings, further confirming the reliability of S100 A12 as a potential biomarker.

Discussion

Strictly speaking, COVID-19 is not a single disease but a clinical syndrome. Studies have shown that because of its association with immunity and inflammation, COVID-19 infection causes more than just pneumonia, especially in elderly patients with underlying diseases such as hypertension and diabetes, and is also associated with

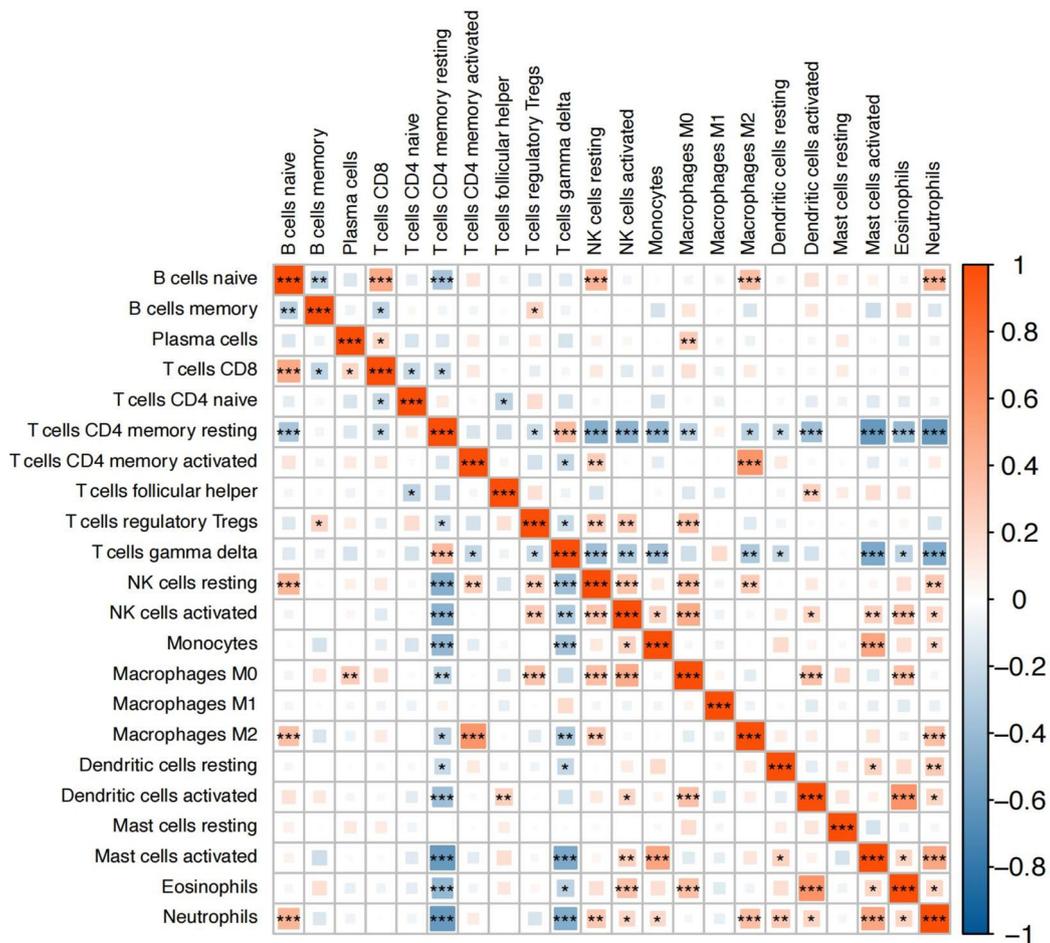


Fig.6 The heatmap matrix is used to demonstrate the correlation between different types of immune cells. The colors and significance markers reflect the strength and significance of the correlation between the cells. Red: positive correlation, the darker the color, the stronger the correlation. Blue: negative correlation, the darker the color, the stronger the negative correlation. White: weak or no significant correlation. P values were showed as: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$

thrombosis in the organism, which can lead to dreadful diseases such as AMI and AIS. Consequently, it is crucial to investigate new targets and biomarkers for the early diagnosis of thrombotic diseases associated with COVID-19.

Under normal circumstances, the immune system plays a protective role for the body. However, in cases of severe or prolonged inflammatory stimuli, this protective effect can become detrimental, leading to negative outcomes such as a hypercoagulable state in the blood [25, 26]. Numerous reports have indicated a correlation between COVID-19 infection and coagulopathy, including thrombus formation. Early in the infection, certain laboratory markers can reflect changes in the level of inflammation, such as elevated levels of D-dimer and fibrinogen. The underlying causes of coagulopathy may involve factors such as hypoxia, direct viral infection, and the induction of tissue factor production by pro-inflammatory

cytokines, including IL-1, IL-6, and TNF. These cytokines activate the extrinsic coagulation pathway [27–29]. It has been shown that the inflammatory response also attenuates certain anticoagulant mechanisms, such as the antithrombin and proteinase C pathways [30, 31]. In addition, novel coronaviruses can also cause direct damage to vascular endothelial cells, thereby reducing their antithrombotic activity [32, 33]. It has been shown that endothelial dysfunction and endothelial cell damage in COVID-19 patients can cause increased vascular endothelial permeability and increased endothelial cell concentration in the systemic circulation, leading to higher levels of soluble intercellular adhesion molecules, which further exacerbate the hypercoagulable state of the blood, and that the massive release of tissue factor (TF) and neutrophils promotes the formation of immune thrombi [34–36]. More specifically, COVID-19 is a multisystem disease involving endothelial cells from

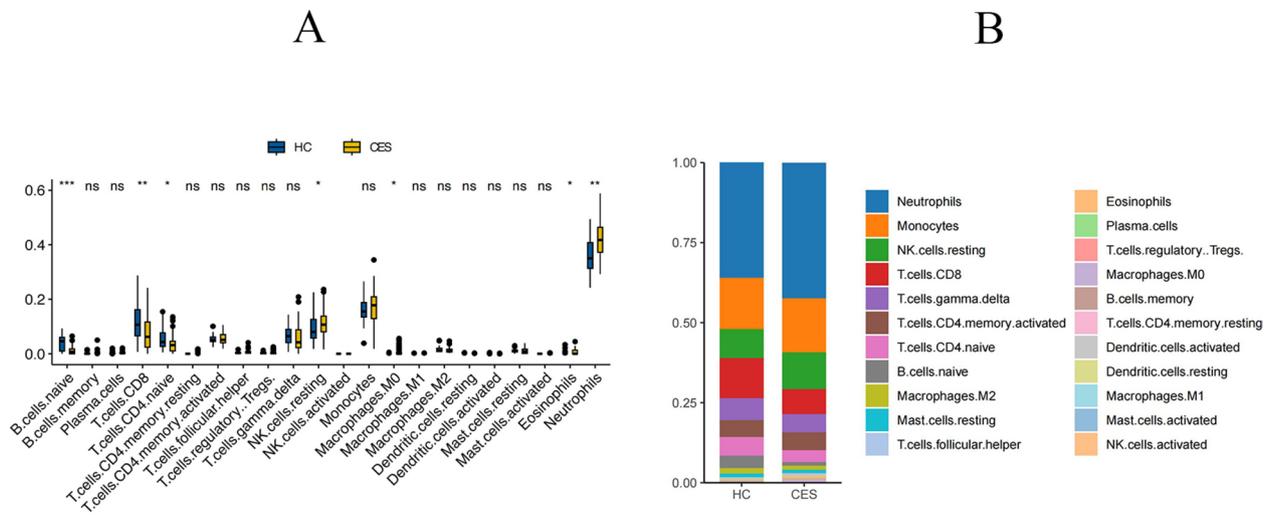


Fig.7 (A) Boxplots were used to show the proportional distribution of different types of immune cells in HC group and AIS(CES) group. The results showed that 7/22 immune cells were significantly different between the AIS and HC groups. (B) Stacked bar graphs were used to demonstrate the immune cell composition of HC group and AIS group. The results showed a difference in immune cell levels between the two groups. P values were showed as:*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$

multiple systems. And the more severe the neocoronary pneumonia infection, the higher the levels of markers of endothelial damage, such as von willebrand factor (VWF), intercellular adhesion molecules including intercellular cell adhesion molecule-1(ICAM-1) and vascular cell adhesion molecule-1(VCAM-1), which can promote thrombosis [25, 37, 38]. Studies have shown a significant increase in the incidence of cardiovascular disease in patients with COVID-19, particularly thrombosis-related diseases such as AMI, AIS and pulmonary embolism, with rates in critically ill patients tens of times higher than in patients with mild infections [3–40].

Research has demonstrated a seven-fold increase in the incidence of myocardial infarction within one week of COVID-19 infection, and a long-term cardiovascular risk after infection, with stroke being the most common manifestation of cerebrovascular disease [41–43]. AMI and AIS are the same as thrombotic diseases, and they are very closely related to each other. After an AMI, certain factors promote thrombus formation, such as post-infarction myocardial ischaemia, dyskinesia and hypercoagulability of the blood. These clots do not only stay in the heart and cause cardiac complications, but also reach the body’s centre through the circulatory system and cause cerebral embolism, leading to AIS. Studies have shown that CES account for a large proportion of all ischaemic strokes, and that the risk of AIS increases dramatically within a short period of time after myocardial infarction [44, 45]. Because of the decline in cardiac ejection volume, left ventricular dilatation and ventricular remodelling occur in long-term conditions and also

increase the risk of cardiogenic stroke [46]. The interaction between AMI and CES suggests that heart disease negatively affects cerebral perfusion, thereby increasing the risk and severity of stroke. Neocoronavirus also crosses the blood–brain barrier and directly attacks brain endothelial cells, inducing hypercoagulability of blood in the brain and promoting thrombosis [7]. Therefore, clinical practice should emphasise early diagnosis, as well as monitoring and management of patients with new crown pneumonia who may develop AMI and AIS. Currently, there are relatively few studies on the correlation between COVID-19, AMI and AIS. We applied bioinformatics methods to explore these three diseases using three GEO datasets, GSE171110, GSE66360, and GSE58294, and found co-DEGs among them. To investigate how the genes affect disease onset and progression at the molecular level, we used GO and KEGG enrichment analyses to make predictions about the possible functions and pathways of these genes. Enrichment analysis showed that co-DEGs were significantly enriched in leukocytes, neutrophils and other related immune-inflammatory responses. The immune system plays a crucial role in the progression of diseases, and neutrophils are intrinsic immune cells, which are activated during infections and play an active role in the body’s immunity, but their over-activation is harmful to the human body, as it can induce excessive immune responses and lead to microvascular and even macrovascular thrombosis, which can lead to inadequate perfusion of organs and tissues, and is one of the main causes of AMI and AIS, and excessive immune responses can directly damage the

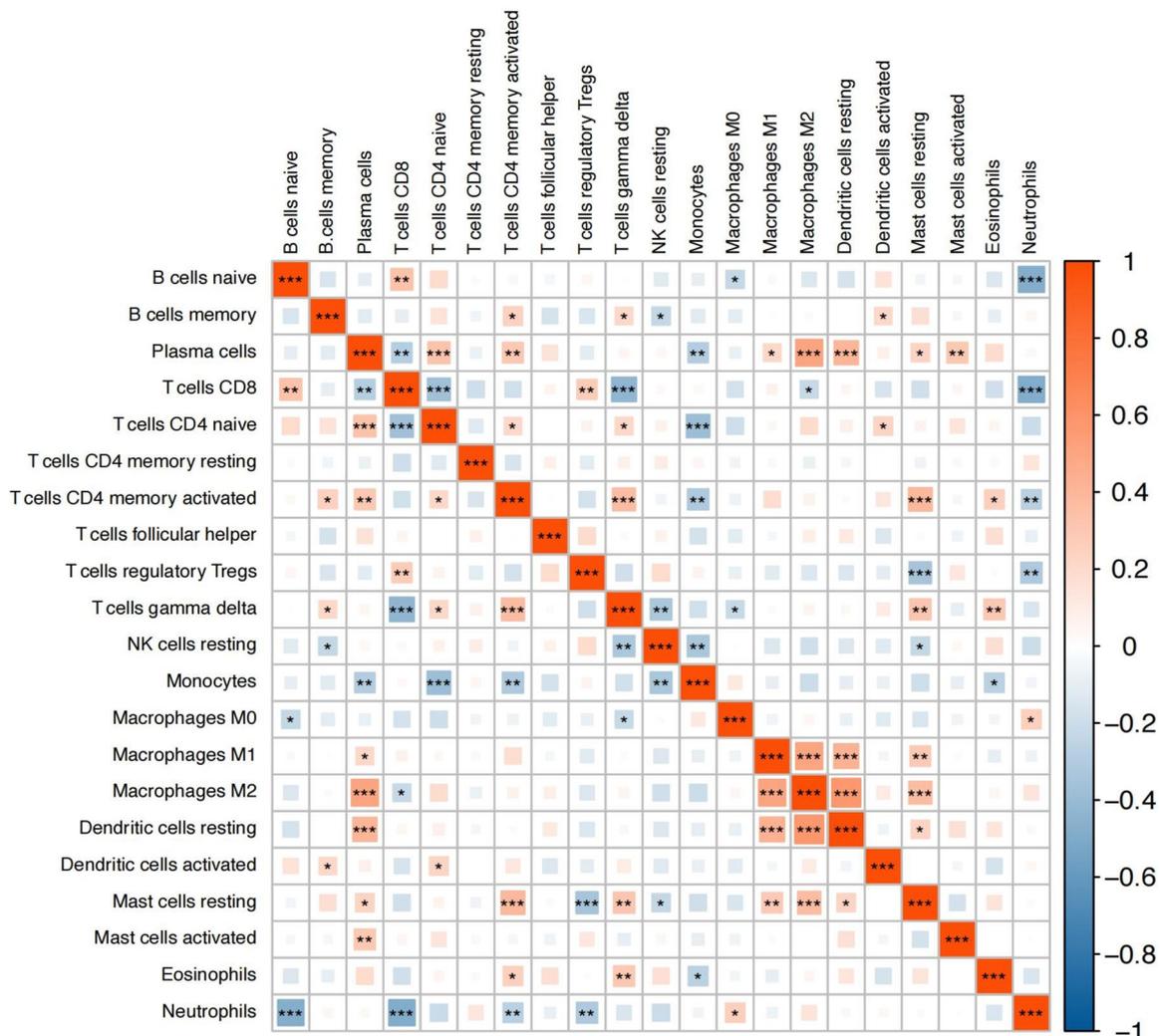


Fig.8 The heatmap matrix is used to demonstrate the correlation between different types of immune cells. The colors and significance markers reflect the strength and significance of the correlation between the cells. Red: positive correlation, the darker the color, the stronger the correlation. Blue: negative correlation, the darker the color, the stronger the negative correlation. White: weak or no significant correlation. P values were showed as: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$

cardio-cerebral vascular system [47]. It has been shown that COVID-19 infection can increase the ability of neutrophils to produce neutrophil extracellular traps (NETs), and NETs have been shown to be a biomarker of the severity of COVID-19 infection [48, 49]. Activated neutrophils can secrete NETs, which consists of a meshwork of DNA, histones and other substances to further kill pathogens. However, its formation may lead to increased inflammation in the organism and promote coagulation, vascular occlusion, and thrombosis [50, 51]. NETs can be degraded by circulating DNA enzymes, and the products of degradation, such as cell-free DNA (cfDNA) and histones, can activate intrinsic pathways of coagulation, which can further contribute to the onset of disease

[52–54]. In addition, COVID-19 leads to an increase in circulating NETs and attacks the body’s endothelial cells, also causing VCAM-1 and ICAM-1 present in endothelial cells to be released, further promoting blood hypercoagulability [55, 56]. Studies have found that in patients who recovered from COVID-19, the neutrophil-to-lymphocyte ratio (NLR) at the time of admission was significantly lower compared to those whose condition worsened. Therefore, NLR can serve as a biomarker for the early prognosis of COVID-19 patients [57]. In COVID-19 patients, lymphocyte reduction is commonly observed, including both CD4 + T cells and CD8 + T cells [58]. It is worth noting that $\gamma\delta$ T cells, a specialized subset of lymphocytes, also exhibit a reduction

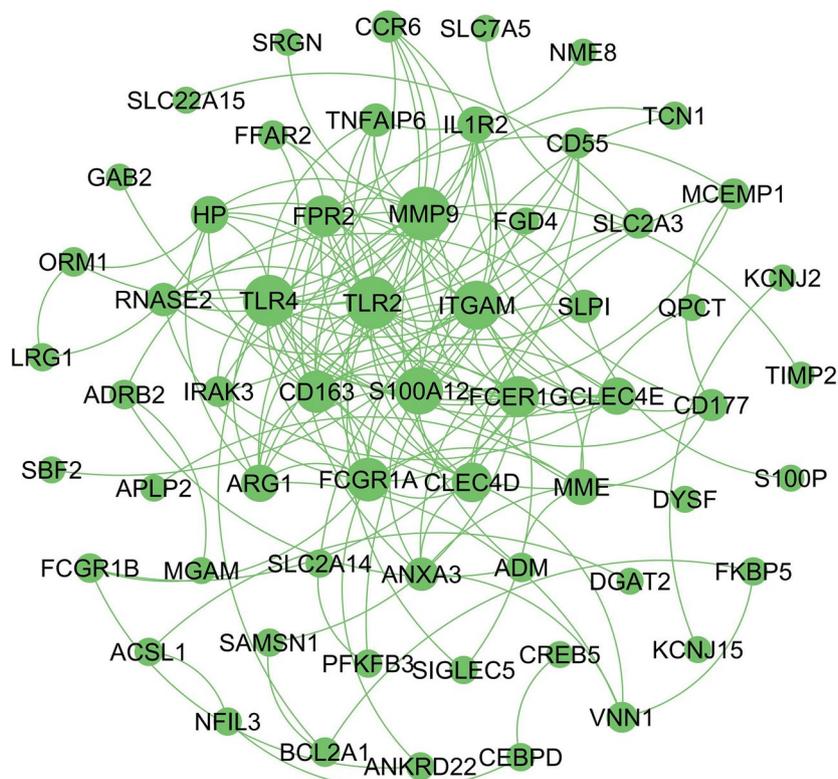


Fig. 9 PPI network and hub genes. Proteins are represented as nodes and functional relationships by edges. The top 10 most influential genes were MMP9, TLR2, TLR4, ITGAM, S100 A12, FCGR1 A, CD163, FCER1G, FPR2, and CLEC4D

in number due to COVID-19 infection [59]. The reduction in $\gamma\delta$ T cells may be related to excessive activation and cellular exhaustion. Certain cytokines, such as IL-10, IL-6, and TNF- α , show a negative correlation with T cell count, suggesting that these cytokines may also play a role in the process of T cell reduction [60, 61]. In addition, changes in spleen function are also involved in the process of reduced B cell and T cell lymphocyte counts [62]. It is noteworthy that mast cells may become a therapeutic target for alleviating the characteristic airway inflammation during the hyperacute phase of COVID-19 [63]. Activated neutrophils and monocytes interact with platelets and the coagulation cascade, leading to intravascular thrombosis in small and large blood vessels [64].

Fibronectin is a substance that plays a crucial role in the recognition and activation of the lectin pathway (LP), one of the complement agglutination pathways, and of which neutrophils are the main source. It is found predominantly in cells, with low levels in the circulation. There are three types of fibronectin, of which ficolin-1 may be involved in part of the body's innate immunity, meaning that it contributes to the body's immune defence against infection [65]. R. Zangari et al. demonstrated that ficolin-1 is heavily depleted within a short

period of time after AIS and can be used as a prognostic predictive marker in patients with AIS [66, 67]. In addition, ficolin-3 has also been reported to be associated with AIS and subarachnoid haemorrhage [68, 69].

The relationship between COVID-19 and the immune mechanisms involved in the pathogenesis of cardiovascular disease is intricate and multifaceted. The results of our enrichment analysis of COVID-19, AMI and AIS co-DEGs showed that co-DEGs were mainly associated with neutrophils, leukocytes and immune responses, indicating that the three diseases mentioned above are extremely related to inflammatory immune responses. Therefore, we later performed immune infiltration analysis, which revealed significant differences in the content of immune cells—including B cells, T cells, NK cells, neutrophils, eosinophils, dendritic cells, and macrophages—between the three diseases and normal samples. After an organism is infected with a virus, immune cells such as neutrophils, eosinophils and NK cells react first and can directly kill the pathogen; this immune response is known as non-specific immunity. In addition, specific immunity will also play a role sometime after the infection occurs, which relies on the antigen presentation of dendritic cells, macrophages and B cells, thus activating the specific immune

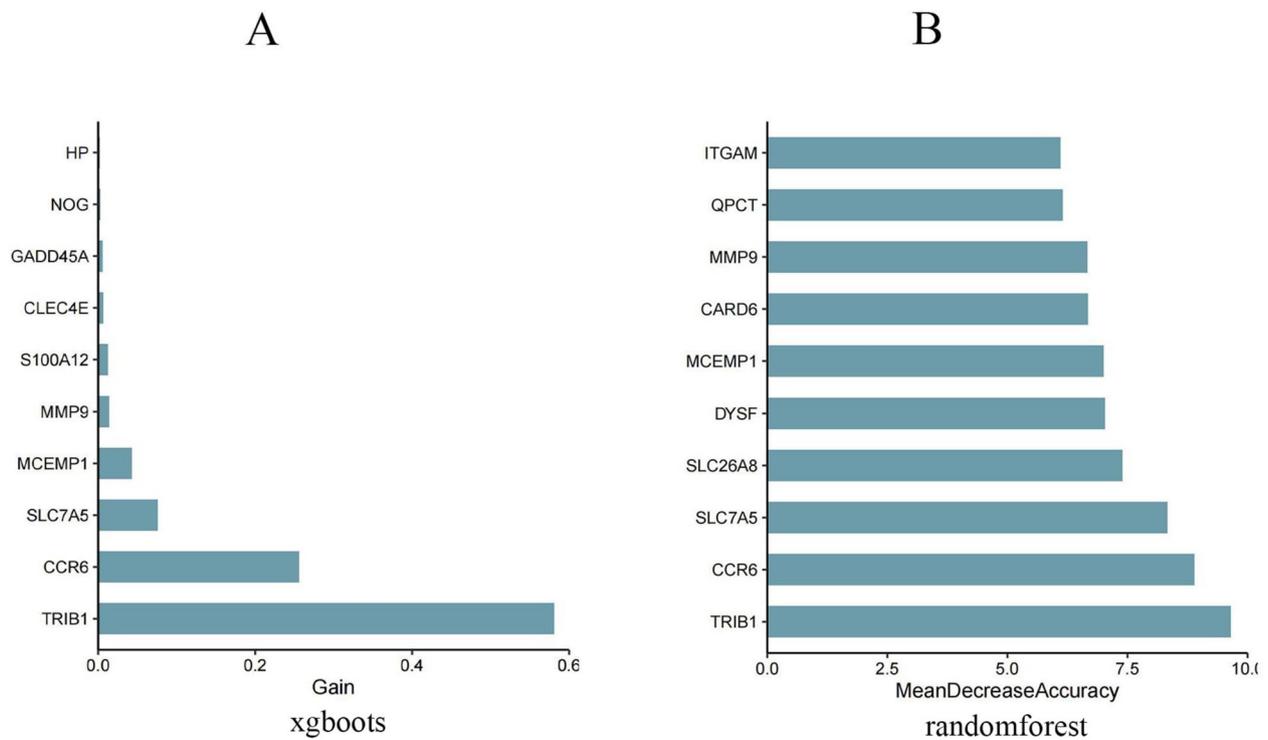


Fig. 10 Feature importance analysis using two machine learning methods: XGBoost and random forest model. **A**) corresponds to COVID-19, **B**) to AML and **C**) to AIS. These plots show the rankings of feature importance derived from the XGBoost model (Plot A) and the Random Forest model (Plot B). In plot A, feature importance is assessed by gain (Gain), which indicates the contribution of each gene to the predictive power of the model. In plot B, importance is measured by mean accuracy decrease, which indicates the importance of each gene to the model's prediction accuracy

response of T cells and B cells [70, 71]. The pathological features of COVID-19 are primarily characterized by significant infiltration of macrophages and neutrophils in the lungs, which subsequently leads to ARDS, including typical manifestations such as alveolar epithelial cell injury, hyaline membrane formation, and interstitial fibrosis [72, 73]. The highly activated pulmonary macrophages, derived from infiltrating inflammatory monocytes, promote the massive release of pro-inflammatory cytokines (such as IL-6 and TNF- α) and the abnormal recruitment of cytotoxic effector cells (such as neutrophils and CD8 + T cells), ultimately leading to further exacerbation of local tissue damage [74]. Lymphocytes play a critical role in antiviral immunity. CD8 + cytotoxic T cells (CTL) directly mediate viral clearance through the release of effector molecules such as perforin, granzymes, and IFN- γ . Additionally, CD4 + helper T cells can synergistically enhance the ability of T cells and B cells to combat viral infections. However, under conditions of chronic viral infection or persistent antigen exposure, T cells may enter an exhausted state, thereby weakening the effectiveness of antiviral immunity [75–78]. As previously mentioned, in COVID-19 patients, NETs released

by activated neutrophils are one of the key mechanisms mediating coagulation dysfunction and the development of thrombotic diseases. Additionally, the inflammatory microenvironment induces the expression of tissue factor by endothelial cells, macrophages, and neutrophils, significantly enhancing the activation of the coagulation cascade in the lungs. The formation of microthrombi may also lead to dysfunction in other organs [64, 79]. The disruption of the blood–brain barrier caused by AIS creates conditions for the migration of immune cells into the brain. White blood cells and other immune cells, including monocytes/macrophages, neutrophils, and lymphocytes from the bloodstream, migrate to the brain and induce inflammatory responses [80–82]. As the earliest subpopulation of leukocytes to infiltrate ischaemic brain tissue, neutrophils persist in the cerebral microvasculature and further exacerbate blood–brain barrier disruption by releasing agents such as protein hydrolases [83]. As the earliest infiltrating T cell subset in ischemic brain tissue, CTL can release effector molecules such as perforin and granzymes after cell–cell interactions and antigen-dependent activation, leading to neuronal death and exacerbating brain tissue damage [84, 85].

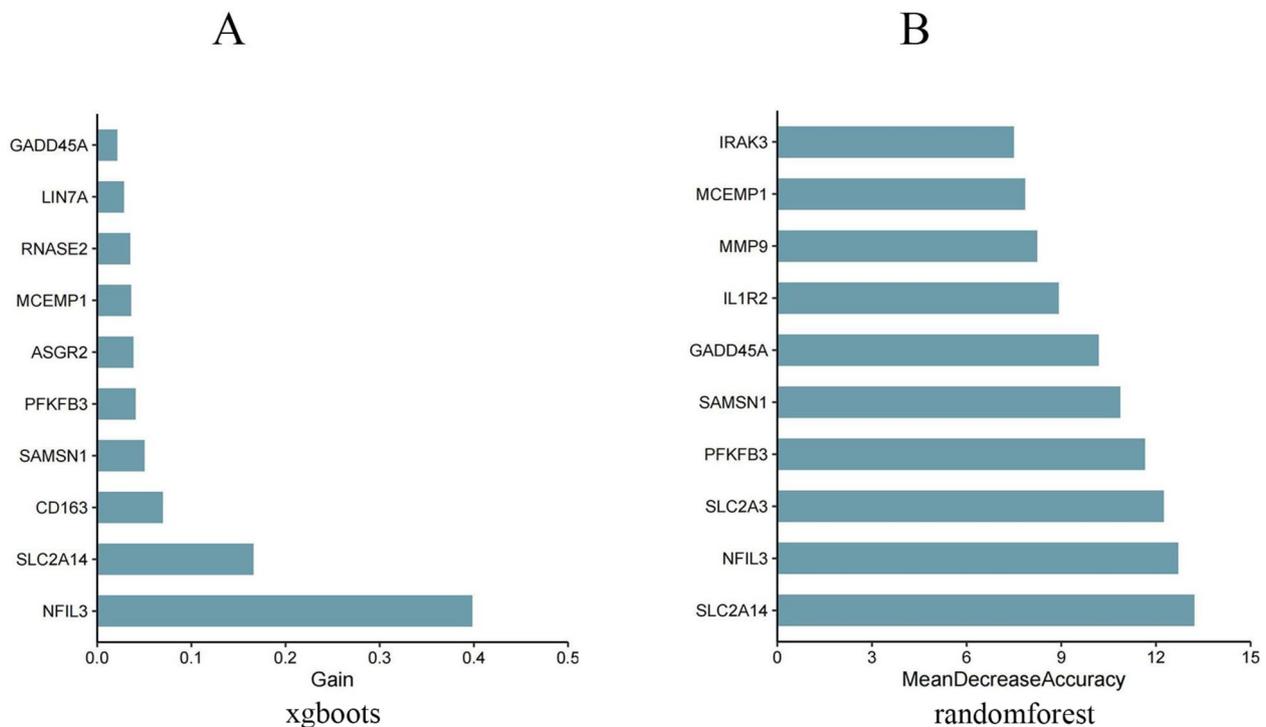


Fig. 11 Feature importance analysis using two machine learning methods: XGBoost and random forest model. **A**) corresponds to COVID-19, **B**) to AMI and **C**) to AIS. These plots show the rankings of feature importance derived from the XGBoost model (Plot A) and the Random Forest model (Plot B). In plot A, feature importance is assessed by gain (Gain), which indicates the contribution of each gene to the predictive power of the model. In plot B, importance is measured by mean accuracy decrease, which indicates the importance of each gene to the model's prediction accuracy

The immune cell response following AMI exhibits a distinct temporal pattern: within 6 to 24 h after infarction, neutrophils are the first to infiltrate the infarcted area, followed by a significant accumulation of pro-inflammatory monocytes and macrophages between 48 to 72 h, collectively mediating myocardial ischemia/reperfusion injury (IRI)-related cardiomyocyte death and tissue damage. Subsequently, the process enters an anti-inflammatory repair phase (days 4–7), which is primarily dominated by anti-inflammatory monocytes/macrophages. These cells regulate the inflammatory response after AMI by suppressing, clearing, and limiting the initial pro-inflammatory cells [86, 87]. Interestingly, researchers have found through animal experiments that mice lacking dendritic cells exhibit worse recovery after AMI, suggesting that dendritic cells may play a protective role in the inflammatory response following AMI [88].

We constructed a PPI network and identified the top 10 DEGs common to COVID-19, AMI, and AIS as the most influential genes, and these pivotal genes include MMP9, TLR2, TLR4, ITGAM, S100 A12, FCGR1 A, CD163, FCER1G, FPR2, and CLEC4D. Random Forest and XGBoost analyses of these two machine learning methods can further identify early diagnostic

biomarkers and potential therapeutic targets common to these three diseases. These DEGs are: CLEC4E, S100 A12, IL1R2. The final gene derived from the combined PPI network and machine learning approach is S100 A12, which may be the common early diagnostic biomarker and potential therapeutic target for these three diseases. The C-type lectin domain family 4 member E (CLEC4E) holds significant importance in the field of sterile inflammation. It is typically expressed by leukocytes and activated through inflammatory responses. Denise Veltman and colleagues demonstrated, through complementary experimental methods in pig, mouse, and human samples, the substantial potential of CLEC4E as a biomarker for the severity of AMI [89]. In addition, studies have shown that CLEC4E, along with other genes, can serve as biomarkers for AIS [90]. IL1R2 is a member of the interleukin-1 receptor family (ILRs). It is released in a soluble form (sIL1R2) and exerts a negative regulatory effect on the IL-1 system, primarily participating in the regulation of local inflammation [91]. According to relevant reports, IL1R2 is highly expressed in AIS patients and is closely associated with the severity of coronary artery disease. Additionally, IL1R2 serves as a biomarker for AMI, and

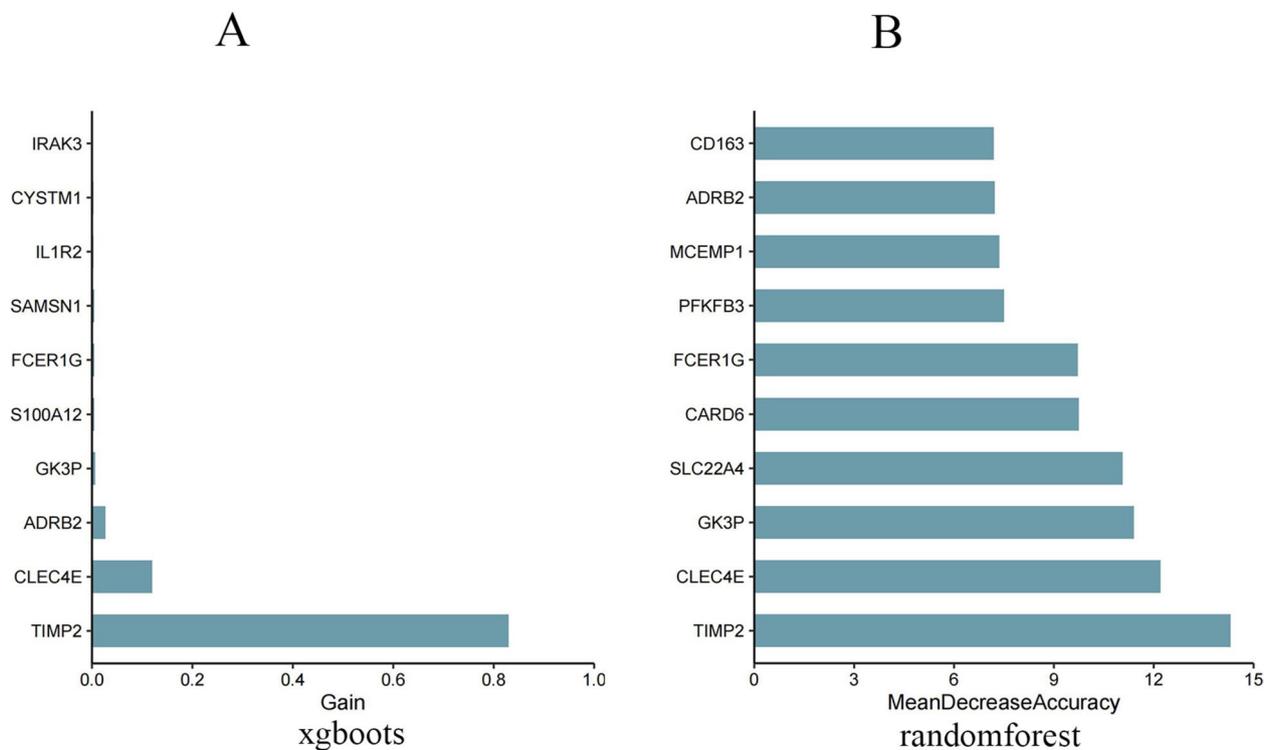


Fig. 12 Feature importance analysis using two machine learning methods: XGBoost and random forest model. **A**) corresponds to COVID-19, **B**) to AMI and **C**) to AIS. These plots show the rankings of feature importance derived from the XGBoost model (Plot A) and the Random Forest model (Plot B). In plot A, feature importance is assessed by gain (Gain), which indicates the contribution of each gene to the predictive power of the model. In plot B, importance is measured by mean accuracy decrease, which indicates the importance of each gene to the model's prediction accuracy

plays a role in the development of pulmonary fibrosis in severe COVID-19 patients [92–95].

S100 A12, identified as a biomarker for the three diseases through various methods, is a calcium-binding protein expressed by neutrophils and monocyte macrophages and belongs to the S100 protein subfamily of myeloid-associated proteins [96, 97]. It plays a crucial role in the body's defense against inflammation and in maintaining immune homeostasis. During the pathological process of diseases, its expression significantly increases, making it a potential biomarker for disease diagnosis and early intervention. Numerous studies have reported that S100 A12 promotes the secretion of pro-inflammatory cytokines and is associated with the development of many diseases, such as coronary artery disease, ARDS, and cerebrovascular disease [96, 98–103]. In addition, it has been demonstrated that S100 A12 is a hub gene immunologically related to AMI and AIS [103, 104]. Interestingly, S100 A12 was found to be associated with the severity of COVID-19 [105]. It has been reported that the receptor for advanced glycation end products (RAGE) acts as a specific receptor for the S100 A12 protein, primarily located on the surface of endothelial cells,

monocytes, and macrophages. S100 A12 can bind to these cell surface receptors and activate them. This interaction promotes the expression of ICAM-1 and VCAM-1 on endothelial cells [98, 106]. This may be related to the widespread endothelial cell damage across multiple systems caused by COVID-19, which further exacerbates the hypercoagulable state in the body, leading to a sharp increase in the incidence of thrombotic-related diseases, such as AMI and AIS. The mechanism by which S100 A12 contributes to the development of thrombotic diseases may involve multiple aspects. Firstly, S100 A12 binds with high affinity to receptor for RAGE, activating the NF- κ B signaling pathway, which in turn upregulates the expression levels of related cytokines. Additionally, S100 A12 may participate in the pathological process of thrombosis through its chemotactic activity. Researchers constructed a transgenic mouse model with smooth muscle cell-specific expression of S100 A12 and found that this model exhibited significantly increased calcification in the coronary arteries and aorta, along with reduced stability of arterial plaques [97, 106, 107]. Patricia Mester et al. found in their study of serum from moderate and severe COVID-19 patients that the severity of COVID-19

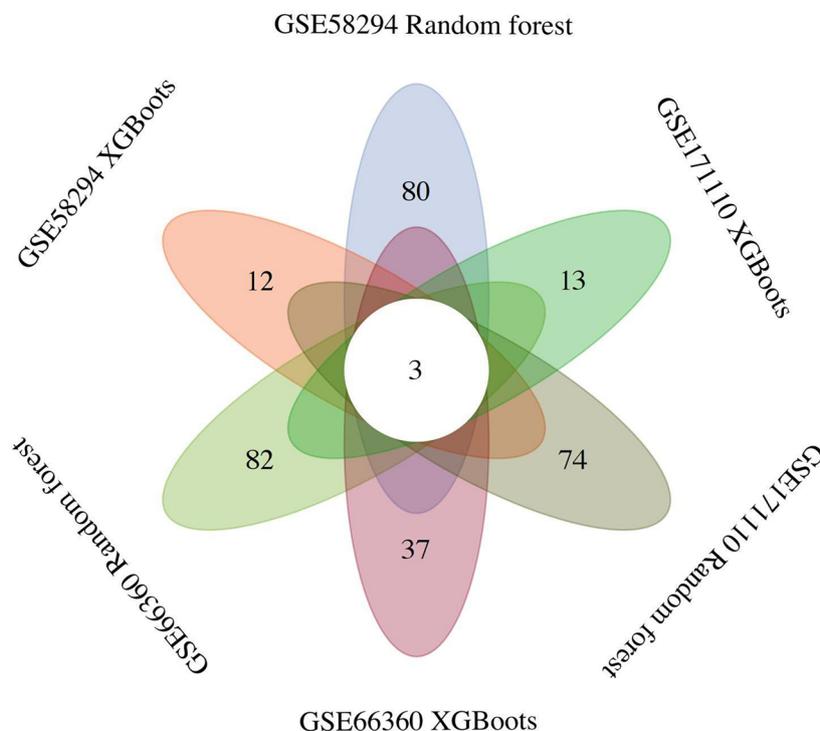


Fig.13 The Venn diagram is used to show the overlap of important genes obtained from different datasets when feature screening is performed using XGBoost and Random Forest models. The different ellipses in the figure represent the set of feature genes for different combinations of datasets and algorithms, and the overlap in the central region indicates the common genes identified in all combinations. By screening the genes using two separate machine learning methods on each disease dataset, the final co-DEGs for them were CLEC4E, S100 A12, IL1R2

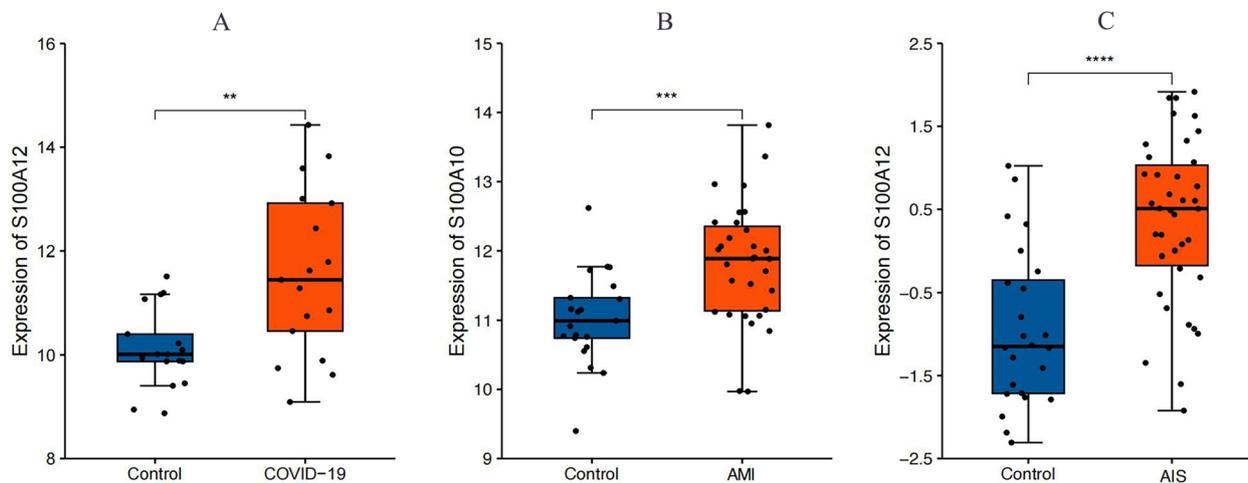


Fig.14 Expression of S100 A12 in the validation dataset. **A** The expression of S100 A12 was increased and significant in the COVID-19 group relative to the HC group. **B** The expression of S100 A12 was increased and significant in the AML group relative to the HC group. **C** The expression of S100 A12 was increased and meaningful in the AIS group relative to the HC group. Orange indicates the disease group and blue indicates the normal group, *P* values are shown as *, *P* < 0.05; **, *P* < 0.01; ***, *P* < 0.001

was positively correlated with the degree of neutrophil dysfunction. The S100 A12 levels in severe COVID-19 patients were higher than in moderate patients, but

the increase in neutrophil count was less pronounced than that of S100 A12, indicating that neutrophil function was impaired in severe COVID-19 patients. This

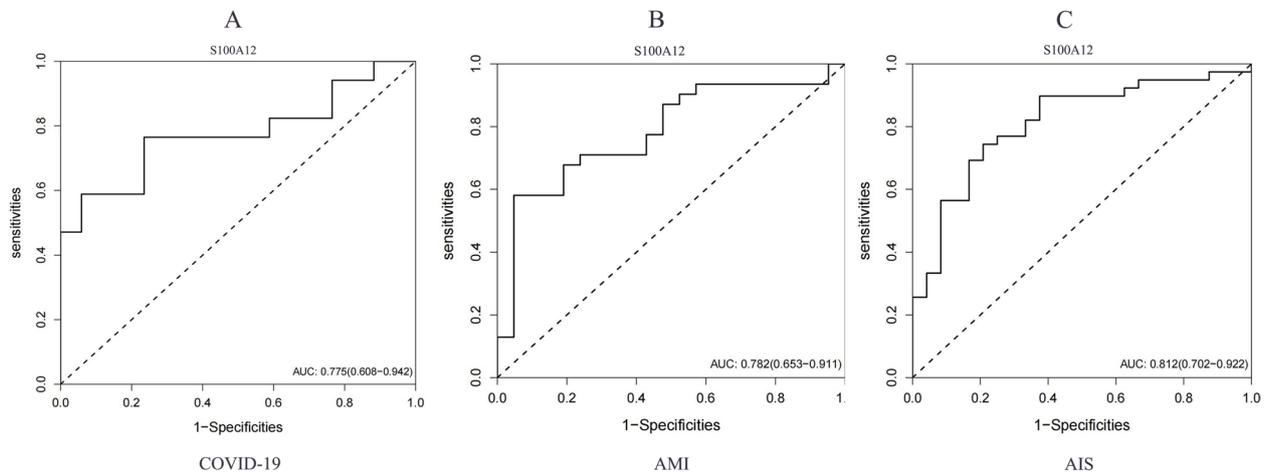


Fig. 15 ROC curves for S100 A12 in the validation dataset. **(A)** ROC curve of S100 A12 in the COVID-19 dataset. **(B)** ROC curve of S100 A12 in the AMI dataset. **(C)** ROC curve of S100 A12 in the AIS dataset. The horizontal coordinate is the rate of false positives, expressed as 1-specificity, and the vertical coordinate is the rate of true positives, expressed as sensitivity

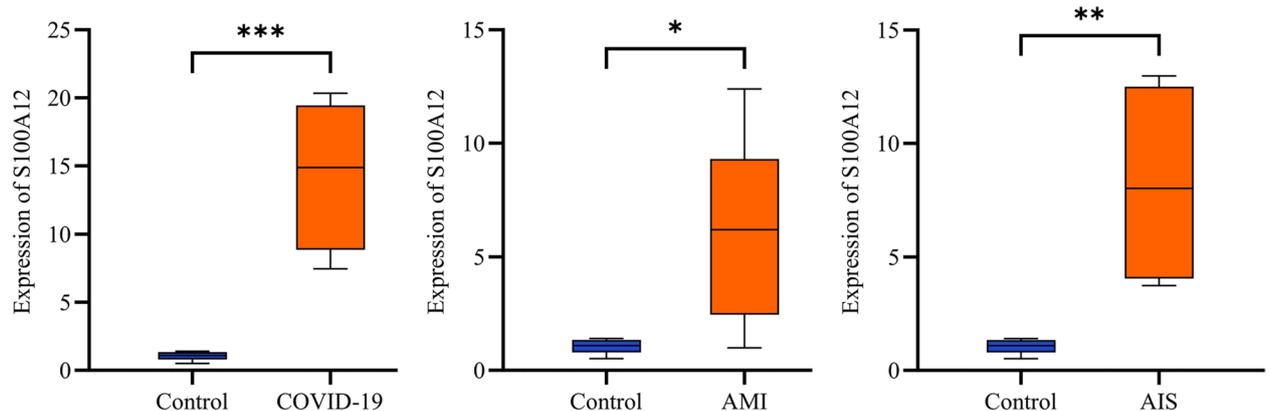


Fig. 16 QRT-PCR was performed to validate the expression of S100 A12 in patients with COVID-19, AMI, and AIS. S100 A12 expression was relatively low in healthy individuals, while significant up-regulation was observed in patients. Orange indicates the disease group and blue indicates the normal group, *P* values are shown as *, *P* < 0.05; **, *P* < 0.01; ***, *P* < 0.001

study proposed S100 A12 as a marker for the severity of COVID-19 [108]. Furthermore, the lung tissue is rich in RAGE, and elevated levels of S100 A12 in COVID-19 patients bind to RAGE in the lungs, leading to the activation of pulmonary endothelial cells, leukocyte extravasation, and neutrophil accumulation, thereby contributing to the occurrence of acute lung injury and acute respiratory distress syndrome [100, 106].

Although the content and results of our study have largely achieved what was originally envisaged, there are still some shortcomings. Firstly, our research data is based on public databases, and there is a lack of validation of the results. Second, the results based on the sequencing of blood samples were not directly sequenced in the organs where the disease occurs, such

as the lungs and the heart. Finally, this study only used genomics-related methods, without comprehensive multi-omics analysis, and the results are slightly less credible. However, based on a large number of studies, our results are basically consistent with theirs. To further enhance the clinical value of the research, future studies should focus on conducting multicenter, large-sample clinical validation of the diagnostic value of S100 A12, evaluating its diagnostic efficacy across different diseases and stages. Simultaneously, the development of quantitative detection kits, along with the establishment of a rapid detection platform, is essential to meet the timeliness requirements of clinical testing. Early diagnosis and timely clinical interventions based on S100 A12 levels can thereby improve patient

outcomes. Additionally, S100 A12 could serve as a routine screening indicator for high-risk populations and be used in combination with other biomarkers to enhance screening specificity. Through these strategies, S100 A12 is expected to become a bridge connecting basic research and clinical practice, providing innovative solutions for the precise diagnosis and treatment of related diseases.

Conclusion

We investigated the diagnostic biomarkers and immune infiltration characteristics common to COVID-19, AMI and AIS using bioinformatics methods and machine learning. Functional and pathway analyses showed that co-DEGs were mainly enriched in immune-inflammatory responses such as leukocytes and neutrophils. Using two machine learning methods and PPI networks, we obtain the common immune-related therapeutic target and biomarker that may ultimately affect these three diseases: S100 A12. In addition, we employed a dual-validation approach, including independent dataset validation and laboratory validation, and found that it was significantly upregulated in all three diseases, consistent with the experimental results. The ROC curve analysis revealed that S100 A12 demonstrated moderate to good diagnostic performance in distinguishing disease patients from healthy controls, with the most prominent performance observed in AIS (AUC = 0.812). Our study suggests that S100 A12, as a potential biomarker, may offer new insights for the early diagnosis of COVID-19, AMI, and AIS. However, this study has some limitations, such as the small sample size of the datasets, lack of laboratory validation, and failure to account for other potential comorbidities in patients. Therefore, future studies should further validate the diagnostic performance of S100 A12 in different populations and explore its clinical application potential.

Acknowledgements

The authors would like to thank the GEO databases for providing data for the study and those authors who uploaded valuable.

Authors' contributions

Y-NM and S-RM drafted the manuscript. LY and JW searched the literature to identify eligible bioinformatics methods. Y-RW, L-JB and LM analyzed the data. Z-HW and Q-QW finally reviewed the first draft. All authors read and approved the final manuscript.

Funding

This study was supported by the General Hospital of Ningxia Medical University [2023] No. 394 2023 New Entry Masters Talent Training Project.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Data retrieved from the GEO database were uploaded in accordance with the guidelines established by the GEO Ethics, Law and Policy Group, so ethical review and informed consent were not required. The research involving human subjects has been approved by the Medical Research Ethics Review Committee of the General Hospital of Ningxia Medical University (Approval No.: KYLL-2025-0941). The study was conducted in compliance with local laws and institutional requirements, and informed consent was obtained from all participants.

Consent for publication

All authors have given consent for publication.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Geriatrics and Specialty Medicine, General Hospital of Ningxia Medical University, Yinchuan, Ningxia, China. ²School of Basic Medical Sciences, Ningxia Medical University, Yinchuan, Ningxia, China. ³Institute of Medical Sciences, General Hospital of Ningxia Medical University, Yinchuan, Ningxia, China. ⁴Diagnosis and Treatment Engineering Technology Research Center of Nervous System Diseases of Ningxia Hui Autonomous Region, Yinchuan, Ningxia, China. ⁵Neurology Center, Ningxia Medical University General Hospital, Yinchuan, Ningxia, China.

Received: 13 November 2024 Accepted: 28 April 2025

Published online: 08 May 2025

References

- Katz JM, et al. COVID-19 Severity and Stroke: Correlation of Imaging and Laboratory Markers. *American J Of Neuroradiology*. 2020;42(2):257–61. <https://doi.org/10.3174/ajnr.a6920>.
- Merkler AE, et al. Risk of Ischemic Stroke in Patients With Coronavirus Disease 2019 (COVID-19) vs Patients With Influenza. *JAMA Neurol*. 2020;77(11):1366. <https://doi.org/10.1001/jamaneurol.2020.2730>.
- Xie Y, Xu E, Bowe B, Al-Aly Z. Long-term cardiovascular outcomes of COVID-19. *Nat Medicine*. 2022;28(3):583–90. <https://doi.org/10.1038/s41591-022-01689-3>.
- Bull TM. Clotting and COVID-19. *Chest*. 2021;159(6):2151–2. <https://doi.org/10.1016/j.chest.2021.02.067>.
- H. Jing, X. Wu, M. Xiang, L. Liu, V. A. Novakovic, and J. Shi, "Pathophysiological mechanisms of thrombosis in acute and long COVID-19" *Frontiers in Immunology*, vol. 13, 2022, <https://doi.org/10.3389/fimmu.2022.992384>.
- W. Wang, C.-Y. Wang, S.-I. Wang, and J. C.-C. Wei, "Long-term cardiovascular outcomes in COVID-19 survivors among non-vaccinated population: A retrospective cohort study from the TriNetX US collaborative networks," *eClinicalMedicine*, vol. 53, p. 101619, 2022, <https://doi.org/10.1016/j.eclinm.2022.101619>.
- S. Zhang et al., "COVID-19 and ischemic stroke: Mechanisms of hypercoagulability (Review)," *International J. Of Molecular Medicine*, vol. 47, no. 3, 2021, <https://doi.org/10.3892/ijmm.2021.4854>.
- Martín-Rojas RM, et al. COVID-19 coagulopathy: An in-depth analysis of the coagulation system. *European J Of Haematology*. 2020;105(6):741–50. <https://doi.org/10.1111/ejh.13501>.
- M. E. Ritchie et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," *N.a. Acids Research*, vol. 43, no. 7, pp. e47–e47, 2015, <https://doi.org/10.1093/nar/gkv007>.
- K. Ito and D. Murphy, "Application of ggplot2 to Pharmacometric Graphics" *CPT: Pharmacometrics & Systems Pharmacology*, vol. 2, no. 10, p. 79, 2013, <https://doi.org/10.1038/psp.2013.56>.
- F. Lam et al., "VennDiagramWeb: a web application for the generation of highly customizable Venn and Euler diagrams," *Bmc Bioinform.*, vol. 17, no. 1, 2016, <https://doi.org/10.1186/s12859-016-1281-5>.

12. A. Ali, A. Ajil, A. M. Sundaram, and N. Joseph, "Detection of Gene Ontology Clusters Using Biclustering Algorithms," *SN Computer Science*, vol. 4, no. 3, 2023, <https://doi.org/10.1007/s42979-022-01624-w>.
13. Ali A, Hulipal VR, Patil SS, Abdulkader R. DPBic: detecting essential proteins in gene expressions using encoding and biclustering algorithm. *J Of Ambient Intelligence And Humanized Computing*. 2021. <https://doi.org/10.1007/s12652-021-03036-9>.
14. C. Chen et al., "Comprehensive single-cell transcriptomic and proteomic analysis reveals NK cell exhaustion and unique tumor cell evolutionary trajectory in non-keratinizing nasopharyngeal carcinoma," *J. Of Translational Medicine*, vol. 21, no. 1, 2023, <https://doi.org/10.1186/s12967-023-04112-8>.
15. Newman AM, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–7. <https://doi.org/10.1038/nmeth.3337>.
16. X. Zhou et al., "Using random forest and biomarkers for differentiating COVID-19 and Mycoplasma pneumoniae infections," *Scientific Reports*, vol. 14, no. 1, 2024, <https://doi.org/10.1038/s41598-024-74057-5>.
17. J. Guo, Q. He, and Y. Li, "Machine learning-based prediction of vitamin D deficiency: NHANES 2001–2018," *Frontiers in Endocrinology*, vol. 15, 2024, <https://doi.org/10.3389/fendo.2024.1327058>.
18. A. Ali, H. V. Ramachandra, N. Ramakrishnan, and G. Murtugudde, "Prognosis of Chronic Kidney Disease Using ML Optimization Techniques," in *2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECCIS)*, 2023, pp. 459–463. <https://doi.org/10.1109/icaecis58353.2023.10170077>.
19. H. V. Ramachandra, A. Ali, P. S. Ambili, S. Thota, and P. N. Asha, "An Optimization on Bicluster Algorithm for Gene Expression Data," in *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)*, 2023, pp. 1–6. <https://doi.org/10.1109/gcat59970.2023.10353373>.
20. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res*. 2018;18(2):623–32. <https://doi.org/10.1021/acs.jproteome.8b00702>.
21. Dykes J, Brunsdon C. Geographically Weighted Visualization: Interactive Graphics for Scale-Varying Exploratory Analysis. *IEEE Trans Visual Comput Graphics*. 2007;13(6):1161–8. <https://doi.org/10.1109/tvcg.2007.70558>.
22. Q. Cheng, X. Chen, H. Wu, and Y. Du, "Three hematologic/immune system-specific expressed genes are considered as the potential biomarkers for the diagnosis of early rheumatoid arthritis through bioinformatics analysis," *J. Of Translational Medicine*, vol. 19, no. 1, 2021, <https://doi.org/10.1186/s12967-020-02689-y>.
23. Ding CG, Jane T-D. Using SAS PROC CALIS to fit Level-1 error covariance structures of latent growth models. *Behav Res Methods*. 2012;44(3):765–87. <https://doi.org/10.3758/s13428-011-0171-z>.
24. Wen T, Niu G, Chen T, Shen Q, Yuan J, Liu Y-X. The best practice for microbiome analysis using R. *Protein Cell*. 2023;14(10):713–25. <https://doi.org/10.1093/procel/pwad024>.
25. Turner S, Khan MA, Putrino D, Woodcock A, Kell DB, Pretorius E. Long COVID: pathophysiological factors and abnormalities of coagulation. *Trends Endocrinol Metab*. 2023;34(6):321–44. <https://doi.org/10.1016/j.tem.2023.03.002>.
26. Hanff TC, Mohareb AM, Giri J, Cohen JB, Chirinos JA. Thrombosis in <scp>COVID</scp>-19. *American J Of Hematology*. 2020;95(12):1578–89. <https://doi.org/10.1002/ajh.25982>.
27. Sriram K, Insel PA. Inflammation and thrombosis in COVID-19 pathophysiology: proteinase-activated and purinergic receptors as drivers and candidate therapeutic targets. *Physiol Rev*. 2021;101(2):545–67. <https://doi.org/10.1152/physrev.00035.2020>.
28. Ackermann M, et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *New England J Of Medicine*. 2020;383(2):120–8. <https://doi.org/10.1056/nejmoa2015432>.
29. J. Helms et al., "High risk of thrombosis in patients with severe SARS-CoV-2 infection: a multicenter prospective cohort study," *N.a. Care Medicine*, vol. 46, no. 6, pp. 1089–1098, 2020, <https://doi.org/10.1007/s00134-020-06062-x>.
30. M. Levi and T. van der Poll, "Two-Way Interactions Between Inflammation and Coagulation," *N.a. In Cardiovascular Medicine*, vol. 15, no. 7, pp. 254–259, 2005, <https://doi.org/10.1016/j.tcm.2005.07.004>.
31. Esmon CT. The impact of the inflammatory response on coagulation. *Thromb Res*. 2004;114(5–6):321–7. <https://doi.org/10.1016/j.thromres.2004.06.028>.
32. Teuwen L-A, Geldhof V, Pasut A, Carmeliet P. COVID-19: the vasculature unleashed. *Nat Reviews Immunology*. 2020;20(7):389–91. <https://doi.org/10.1038/s41577-020-0343-0>.
33. Geijteman ECT, Bosch FH, Smulders YM. Universal Do-Not-Resuscitate Orders, Social Worth, and Life-Years. *Annals Of Intern Medicine*. 2020;173(12):1028–9. <https://doi.org/10.7326/120-1196>.
34. Falcinelli E, et al. Role of endothelial dysfunction in the thrombotic complications of COVID-19 patients. *J Of Infection*. 2021;82(5):186–230. <https://doi.org/10.1016/j.jinf.2020.11.041>.
35. J. W. Yau, H. Teoh, and S. Verma, "Endothelial cell control of thrombosis," *BMC Cardiovascular Disorders*, vol. 15, no. 1, 2015, <https://doi.org/10.1186/s12872-015-0124-z>.
36. Gąsecka A, et al. Thrombotic Complications in Patients with COVID-19: Pathophysiological Mechanisms, Diagnosis, and Treatment. *Cardiovasc Drugs Ther*. 2020;35(2):215–29. <https://doi.org/10.1007/s10557-020-07084-9>.
37. Kijpers MJE, Heemskerk JWM, Jurk K. Molecular Mechanisms of Hemostasis, Thrombosis and Thrombo-Inflammation. *International J Of Molecular Sciences*. 2022;23(10):5825. <https://doi.org/10.3390/ijms23105825>.
38. Birnhuber A, et al. Between inflammation and thrombosis: endothelial cells in COVID-19. *European Respiratory J*. 2021;58(3):2100377. <https://doi.org/10.1183/13993003.00377-2021>.
39. Yu J-W, Sun L-J, Zhao Y-H, Kang P, Yan B-Z. Impact of sex on virologic response rates in genotype 1 chronic hepatitis C patients with peginterferon alpha-2a and ribavirin treatment. *International J Of Infectious Diseases*. 2011;15(11):e740–6. <https://doi.org/10.1016/j.ijid.2011.05.018>.
40. Bangalore S, et al. ST-Segment Elevation in Patients with Covid-19 — A Case Series. *New England J Of Medicine*. 2020;382(25):2478–80. <https://doi.org/10.1056/nejmc2009020>.
41. Ho FK, et al. Thromboembolic Risk in Hospitalized and Nonhospitalized COVID-19 Patients. *Mayo Clinic Proc*. 2021;96(10):2587–97. <https://doi.org/10.1016/j.mayocp.2021.07.002>.
42. R. Chang, T. Y.-T. Chen, S.-I. Wang, Y.-M. Hung, H.-Y. Chen, and C.-C. J. Wei, "Risk of autoimmune diseases in patients with COVID-19: a retrospective cohort study," *eClinicalMedicine*, vol. 56, p. 101783, 2023, <https://doi.org/10.1016/j.eclinm.2022.101783>.
43. S. Chen et al., "Prior COVID-19 vaccination and reduced risk of cerebrovascular diseases among COVID-19 survivors," *J. Of Med. Virology*, vol. 96, no. 5, 2024, <https://doi.org/10.1002/jmv.29648>.
44. C. Bogiatzi, D. G. Hackam, A. I. McLeod, and J. D. Spence, "Secular Trends in Ischemic Stroke Subtypes and Stroke Risk Factors," *N.a.*, vol. 45, no. 11, pp. 3208–3213, 2014, <https://doi.org/10.1161/strokeaha.114.006536>.
45. A. E. Merkle et al., "Duration of Heightened Ischemic Stroke Risk After Acute Myocardial Infarction," *J. Of The American N.a. Association*, vol. 7, no. 22, 2018, <https://doi.org/10.1161/jaha.118.010782>.
46. Michaud GF, Stevenson WG. Atrial Fibrillation. *New England J Of Medicine*. 2021;384(4):353–61. <https://doi.org/10.1056/nejmcp2023658>.
47. Stiel L, Meziani F, Helms J. Neutrophil Activation During Septic Shock. *Shock*. 2018;49(4):371–84. <https://doi.org/10.1097/shk.0000000000000980>.
48. F. P. Veras et al., "SARS-CoV-2-triggered neutrophil extracellular traps mediate COVID-19 pathology," *J. Of Experimental Medicine*, vol. 217, no. 12, 2020, <https://doi.org/10.1084/jem.20201129>.
49. Silva V, Radic M. COVID-19 Pathology Sheds Further Light on Balance between Neutrophil Proteases and Their Inhibitors. *Biomolecules*. 2022;13(1):82. <https://doi.org/10.3390/biom13010082>.
50. S.-W. Kim, H. Lee, H.-K. Lee, I.-D. Kim, and J.-K. Lee, "Neutrophil extracellular trap induced by HMGB1 exacerbates damages in the ischemic brain," *Acta Neuropathologica Communications*, vol. 7, no. 1, 2019, <https://doi.org/10.1186/s40478-019-0747-x>.
51. Ravindran M, Khan MA, Palaniyar N. Neutrophil Extracellular Trap Formation: Physiology, Pathology, and Pharmacology. *Biomolecules*. 2019;9(8):365. <https://doi.org/10.3390/biom9080365>.
52. K. Gollomp et al., "Neutrophil accumulation and NET release contribute to thrombosis in HIT," *Jci N.a.*, vol. 3, no. 18, 2018, <https://doi.org/10.1172/jci.insight.99445>.

53. Gould TJ, Lysov Z, Liaw PC. Extracellular DNA and histones: double-edged swords in immunothrombosis. *J Of Thrombosis And Haemostasis*. 2015;13:S82–91. <https://doi.org/10.1111/jth.12977>.
54. A. T. P. Ngo et al., "Platelet factor 4 limits neutrophil extracellular trap- and cell-free DNA-induced thrombogenicity and endothelial injury," *Jci N.a.*, vol. 8, no. 22, 2023, <https://doi.org/10.1172/jci.insight.171054>.
55. Folco EJ, et al. Neutrophil Extracellular Traps Induce Endothelial Cell Activation and Tissue Factor Production Through Interleukin-1 α and Cathepsin G. *Arterioscler Thromb Vasc Biol*. 2018;38(8):1901–12. <https://doi.org/10.1161/atvbaha.118.311150>.
56. Amalakuhan B, et al. Endothelial adhesion molecules and multiple organ failure in patients with severe sepsis. *Cytokine*. 2016;88:267–73. <https://doi.org/10.1016/j.cyto.2016.08.028>.
57. R. Chen et al., "Longitudinal hematologic and immunologic variations associated with the progression of COVID-19 patients in China," *J. Of N.a. And Clinical Immunology*, vol. 146, no. 1, pp. 89–100, 2020, <https://doi.org/10.1016/j.jaci.2020.05.003>.
58. W. Wen et al., "Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing," *Cell Discovery*, vol. 6, no. 1, 2020, <https://doi.org/10.1038/s41421-020-0168-9>.
59. Rijkers G, Vervenne T, van der Pol P. More bricks in the wall against SARS-CoV-2 infection: involvement of γ 9 δ 2 T cells. *Cell Mol Immunol*. 2020;17(7):771–2. <https://doi.org/10.1038/s41423-020-0473-0>.
60. G. von Massow, S. Oh, A. Lam, and K. Gustafsson, "Gamma Delta T Cells and Their Involvement in COVID-19 Virus Infections," *Frontiers in Immunology*, vol. 12, 2021, <https://doi.org/10.3389/fimmu.2021.741218>.
61. B. Diao et al., "Reduction and Functional Exhaustion of T Cells in Patients With Coronavirus Disease 2019 (COVID-19)," *Frontiers in Immunology*, vol. 11, 2020, <https://doi.org/10.3389/fimmu.2020.00827>.
62. C. M. Rossi, M. V. Lenti, S. Merli, and A. D. Sabatino, "Role of IgM Memory B Cells and Spleen Function in COVID-19," *Frontiers in Immunology*, vol. 13, 2022, <https://doi.org/10.3389/fimmu.2022.889876>.
63. Murdaca G, et al. Basophils and Mast Cells in COVID-19 Pathogenesis. *Cells*. 2021;10(10):2754. <https://doi.org/10.3390/cells10102754>.
64. Bonaventura A, et al. Endothelial dysfunction and immunothrombosis as key pathogenic mechanisms in COVID-19. *Nat Reviews Immunology*. 2021;21(5):319–29. <https://doi.org/10.1038/s41577-021-00536-9>.
65. Honoré C, et al. The innate pattern recognition molecule Ficolin-1 is secreted by monocytes/macrophages and is circulating in human plasma. *Mol Immunol*. 2008;45(10):2782–9. <https://doi.org/10.1016/j.molimm.2008.02.005>.
66. R. Zangari et al., "Early ficolin-1 is a sensitive prognostic marker for functional outcome in ischemic stroke," *J. Of Neuroinflammation*, vol. 13, no. 1, 2016, <https://doi.org/10.1186/s12974-016-0481-2>.
67. Kitajima H, Sonoda M, Yamamoto K. HLA and SNP haplotype mapping in the Japanese population. *Genes Immun*. 2012;13(7):543–8. <https://doi.org/10.1038/gene.2012.35>.
68. Füst G, et al. Low ficolin-3 levels in early follow-up serum samples are associated with the severity and unfavorable outcome of acute ischemic stroke. *J Of Neuroinflammation*. 2011;8(1):185. <https://doi.org/10.1186/1742-2094-8-185>.
69. Zanier ER, et al. Ficolin-3-mediated lectin complement pathway activation in patients with subarachnoid hemorrhage. *Neurology*. 2014;82(2):126–34. <https://doi.org/10.1212/wnl.000000000000020>.
70. D. Primorac et al., "Adaptive Immune Responses and Immunity to SARS-CoV-2," *Frontiers in Immunology*, vol. 13, 2022, <https://doi.org/10.3389/fimmu.2022.848582>.
71. Giles JR, Kashgarian M, Koni PA, Shlomchik MJ. B Cell-Specific MHC Class II Deletion Reveals Multiple Nonredundant Roles for B Cell Antigen Presentation in Murine Lupus. *The J Of Immunology*. 2015;195(6):2571–9. <https://doi.org/10.4049/jimmunol.1500792>.
72. Z. Varga et al., "Endothelial cell infection and endotheliitis in COVID-19," *The N.a.*, vol. 395, no. 10234, pp. 1417–1418, 2020, [https://doi.org/10.1016/s0140-6736\(20\)30937-5](https://doi.org/10.1016/s0140-6736(20)30937-5).
73. Tian S, Hu W, Niu L, Liu H, Xu H, Xiao S-Y. Pulmonary Pathology of Early-Phase 2019 Novel Coronavirus (COVID-19) Pneumonia in Two Patients With Lung Cancer. *J Of Thorac Oncology*. 2020;15(5):700–4. <https://doi.org/10.1016/j.jtho.2020.02.010>.
74. R. Knoll, J. L. Schultze, and J. Schulte-Schrepping, "Monocytes and Macrophages in COVID-19," *Frontiers in Immunology*, vol. 12, 2021, <https://doi.org/10.3389/fimmu.2021.720109>.
75. Mescher MF, et al. Signals required for programming effector and memory development by CD8⁺ T cells. *Immunol Rev*. 2006;211(1):81–92. <https://doi.org/10.1111/j.0105-2896.2006.00382.x>.
76. Zhu J, Yamane H, Paul WE. Differentiation of Effector CD4 T Cell Populations. *Annu Rev Immunol*. 2010;28(1):445–89. <https://doi.org/10.1146/annurev-immunol-030409-101212>.
77. Ng CT, Snell LM, Brooks DG, Oldstone MBA. Networking at the Level of Host Immunity: Immune Cell Interactions during Persistent Viral Infections. *Cell Host Microbe*. 2013;13(6):652–64. <https://doi.org/10.1016/j.chom.2013.05.014>.
78. Fenwick C, et al. T-cell exhaustion in HIV infection. *Immunol Rev*. 2019;292(1):149–63. <https://doi.org/10.1111/imr.12823>.
79. Levi M, van der Poll T. Coagulation and sepsis. *Thromb Res*. 2017;149:38–44. <https://doi.org/10.1016/j.thromres.2016.11.007>.
80. Petrovic-Djergovic D, Goonewardena SN, Pinsky DJ. Inflammatory Disequilibrium in Stroke. *Circ Res*. 2016;119(1):142–58. <https://doi.org/10.1161/circresaha.116.308022>.
81. Anrather J, Iadecola C. Inflammation and Stroke: An Overview. *Neurotherapeutics*. 2016;13(4):661–70. <https://doi.org/10.1007/s13311-016-0483-x>.
82. Rayasam A, et al. Immune responses in stroke: how the immune system contributes to damage and healing after stroke and how this knowledge could be translated to better cures? *Immunology*. 2018;154(3):363–76. <https://doi.org/10.1111/imm.12918>.
83. K. Malone, S. Amu, A. C. Moore, and C. Waeber, "Immunomodulatory Therapeutic Strategies in Stroke," *Frontiers in Pharmacology*, vol. 10, 2019, <https://doi.org/10.3389/fphar.2019.00630>.
84. Chu HX, et al. Immune Cell Infiltration in Malignant Middle Cerebral Artery Infarction: Comparison with Transient Cerebral Ischemia. *J Cereb Blood Flow Metab*. 2013;34(3):450–9. <https://doi.org/10.1038/jcbfm.2013.217>.
85. Mracsko E, et al. Antigen Dependently Activated Cluster of Differentiation 8-Positive T Cells Cause Perforin-Mediated Neurotoxicity in Experimental Stroke. *The J Of Neuroscience*. 2014;34(50):16784–95. <https://doi.org/10.1523/jneurosci.1867-14.2014>.
86. Ong S-B, et al. Inflammation following acute myocardial infarction: Multiple players, dynamic roles, and novel therapeutic opportunities. *Pharmacol Ther*. 2018;186:73–87. <https://doi.org/10.1016/j.pharmthera.2018.01.001>.
87. Andreadou I, et al. Immune cells as targets for cardioprotection: new players and novel therapeutic opportunities. *Cardiovasc Res*. 2019;115(7):1117–30. <https://doi.org/10.1093/cvr/cvz050>.
88. T. Nagai et al., "Decreased Myocardial Dendritic Cells is Associated With Impaired Reparative Fibrosis and Development of Cardiac Rupture After Myocardial Infarction in Humans," *J. Of The American N.a. Association*, vol. 3, no. 3, 2014, <https://doi.org/10.1161/jaha.114.000839>.
89. D. Veltman et al., "Clec4e-Receptor Signaling in Myocardial Repair After Ischemia-Reperfusion Injury," *JACC: Basic to Translational Science*, vol. 6, no. 8, pp. 631–646, 2021, <https://doi.org/10.1016/j.jacbs.2021.07.001>.
90. Q. Li et al., "Identification of diagnostic signatures for ischemic stroke by machine learning algorithm," *J. Of N.a. And Cerebrovascular Diseases*, vol. 33, no. 3, p. 107564, 2024, <https://doi.org/10.1016/j.jstrokecerebrovasdis.2024.107564>.
91. Molgora M, Supino D, Mantovani A, Garlanda C. Tuning inflammation and immunity by the negative regulators <sc>IL</sc>-1R2 and <sc>IL</sc>-1R8. *Immunol Rev*. 2017;281(1):233–47. <https://doi.org/10.1111/imr.12609>.
92. Zhao M, Liu A, Wu J, Mo L, Lu F, Wan G. Il1r2 and Tnfrsf12a in transcranial magnetic stimulation effect of ischemic stroke via bioinformatics analysis. *Medicine*. 2024;103(4):e36109. <https://doi.org/10.1097/md.00000000000036109>.
93. Chen Q, Li Z, Wang M, Li G. Over-expression of IL1R2 in PBMCs of Patients with Coronary Artery Disease and Its Clinical Significance. *The Anatol J Of Cardiology*. 2022;26(9):710–6. <https://doi.org/10.5152/anatoljcardiol.2022.1241>.
94. Shi W, et al. Bioinformatics approach to identify the hub gene associated with COVID-19 and idiopathic pulmonary fibrosis. *IET Syst Biol*. 2023. <https://doi.org/10.1049/syb2.12080>.
95. E. Zhao, H. Xie, and Y. Zhang, "Predicting Diagnostic Gene Biomarkers Associated With Immune Infiltration in Patients With Acute Myocardial

- Infarction," *Frontiers in Cardiovascular Medicine*, vol. 7, 2020, <https://doi.org/10.3389/fcvm.2020.586871>.
96. Nazari A, et al. S100A12 in renal and cardiovascular diseases. *Life Sci*. 2017;191:253–8. <https://doi.org/10.1016/j.lfs.2017.10.036>.
 97. Xia P, Ji X, Yan L, Lian S, Chen Z, Luo Y. Roles of S100A8, S100A9 and S100A12 in infection, inflammation and immunity. *Immunology*. 2023;171(3):365–76. <https://doi.org/10.1111/imm.13722>.
 98. Tsirebolos G, et al. Emerging markers of inflammation and oxidative stress as potential predictors of coronary artery disease. *International J Of Cardiology*. 2023;376:127–33. <https://doi.org/10.1016/j.ijcard.2023.02.005>.
 99. Buyukterzi Z, et al. Enhanced S100A9 and S100A12 Expression in Acute Coronary Syndrome. *Biomark Med*. 2017;11(3):229–37. <https://doi.org/10.2217/bmm-2016-0253>.
 100. Wittkowski H, et al. Neutrophil-derived S100A12 in acute lung injury and respiratory distress syndrome. *Crit Care Med*. 2007;35(5):1369–75. <https://doi.org/10.1097/01.ccm.0000262386.32287.29>.
 101. H. Amini et al., "Early peripheral blood gene expression associated with good and poor 90-day ischemic stroke outcomes," *J. Of Neuroinflammation*, vol. 20, no. 1, 2023, <https://doi.org/10.1186/s12974-022-02680-y>.
 102. A. Angerfors et al., "Proteomic profiling identifies novel inflammation-related plasma proteins associated with ischemic stroke outcome," *J. Of Neuroinflammation*, vol. 20, no. 1, 2023, <https://doi.org/10.1186/s12974-023-02912-9>.
 103. Liu W, et al. Identification of biomarkers and immune infiltration in acute myocardial infarction and heart failure by integrated analysis. *Biosci Rep*. 2023;43(7):BSR20222552. <https://doi.org/10.1042/bsr20222552>.
 104. H. Bao, J. Li, B. Zhang, J. Huang, D. Su, and L. Liu, "Integrated bioinformatics and machine-learning screening for immune-related genes in diagnosing non-alcoholic fatty liver disease with ischemic stroke and RRS1 pan-cancer analysis," *Frontiers in Immunology*, vol. 14, 2023, <https://doi.org/10.3389/fimmu.2023.1113634>.
 105. Murphy SL, et al. Circulating markers of extracellular matrix remodelling in severe COVID-19 patients. *J Of Intern Medicine*. 2023;294(6):784–97. <https://doi.org/10.1111/joim.13725>.
 106. Hofmann MA, et al. RAGE Mediates a Novel Proinflammatory Axis. *Cell*. 1999;97(7):889–901. [https://doi.org/10.1016/s0092-8674\(00\)80801-6](https://doi.org/10.1016/s0092-8674(00)80801-6).
 107. Bowman MAH, et al. S100A12 in Vascular Smooth Muscle Accelerates Vascular Calcification in Apolipoprotein E-Null Mice by Activating an Osteogenic Gene Regulatory Program. *Arterioscler Thromb Vasc Biol*. 2011;31(2):337–44. <https://doi.org/10.1161/atvbaha.110.217745>.
 108. Mester P, et al. High Serum S100A12 as a Diagnostic and Prognostic Biomarker for Severity, Multidrug-Resistant Bacteria Superinfection and Herpes Simplex Virus Reactivation in COVID-19. *Viruses*. 2024;16(7):1084. <https://doi.org/10.3390/v16071084>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.