ORIGINAL RESEARCH

A Multi-Machine Learning Consensus Model Based on Clinical Features Reveals That Interleukin-10 Derived from Monocytes Leads to a Poor Prognosis in Patients with Coronavirus Disease-2019

Jing Yu [1], Yike Zhang[1], Zhixiong Ye[1], Kun Tang [1], Yiming Ma[1], Linlin Fu[1], Tongtong Cui[1], Hening Kang[2], Yadong Yuan[1], Wensen Pan [1]

[1]Second Department of Respiratory and Critical Care Medicine, The Second Hospital of Hebei Medical University, Shijiazhuang, People's Republic of China; [2]College of Basic Medicine, Hebei Medical University, Shijiazhuang, People's Republic of China

Correspondence: Wensen Pan, Email 27400010@hebmu.edu.cn

Background: Despite ongoing interventions, SARS-CoV-2 continues to cause significant global morbidity and mortality. Early diagnosis and intervention are crucial for effective clinical management. However, prognostic features based on transcriptional data have shown limited effectiveness, highlighting the need for more precise biomarkers to improve COVID-19 treatment outcomes.

Methods: We retrospectively analyzed 149 clinical features from 189 COVID-19 patients, identifying prognostic features via univariate Cox regression. The cohort was split into training and validation sets, and 77 prognostic models were developed using seven machine learning algorithms. Among these, the least absolute shrinkage and selection operator (Lasso) method was employed to refine the selection of prognostic variables by ten-fold cross-validation strategy, which were then integrated with random survival forests (RSF) to build a robust COVID-19-related prognostic model (CRM). Model accuracy was evaluated across training, validation, and entire cohorts. The diagnostic relevance of interleukin-10 (IL-10) was confirmed in bulk transcriptional data and validated at the single-cell level, where we also examined changes in cellular communication between mononuclear cells with differing IL-10 expression and other immune cells.

Results: Univariate Cox regression identified 43 prognostic features. Among the 77 machine learning models, the combination of Lasso and RSF produced the most robust CRM. This model consistently performed well across training, validation, and entire cohorts. IL-10 emerged as a key prognostic feature within the CRM, validated by single-cell transcriptional data. Transcriptome analysis confirmed the stable diagnostic value of IL-10, with mononuclear cells identified as the primary IL-10 source. Moreover, differential IL-10 expression in these cells was linked to altered cellular communication in the COVID-19 immune microenvironment.

Conclusion: The CRM provides accurate prognostic predictions for COVID-19 patients. Additionally, the study underscores the importance of early IL-10 level testing upon hospital admission, which could inform therapeutic strategies.

Keywords: SARS-CoV-2, machine learning, IL-10, therapeutic guidance, clinical decision making

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the pathogen responsible for the global pandemic of coronavirus disease-2019 (COVID-19), has emerged as a significant threat worldwide.[1,2] It was first reported towards the end of 2019,[3] and since then, it has rapidly spread across more than 150 countries on every continent except Antarctica. Although COVID-19 is relatively less contagious than its predecessor, SARS-CoV-1, it is associated with a higher mortality rate compared to the seasonal flu, estimated to be around 3.4%. The average incubation period of the disease is approximately 5.2 days, and in severe cases, individuals succumb to the illness within a median duration of 14 days after

the onset of symptoms.[4] The leading causes of death in patients with COVID-19 are acute respiratory distress syndrome, hypoxic respiratory failure, and a cytokine storm triggered by a prolonged infection.[5]

While COVID-19 is generally characterized by asymptomatic or mild to moderate symptoms such as fever and dry cough, it can progress into acute pneumonia, particularly in susceptible populations such as the elderly with metabolic, cardiovascular, and pulmonary conditions.[6,7] In terms of diagnosis, RT-PCR testing is the gold standard for early confirmation of COVID-19 infection, with rapid antigen tests developed for large-scale screening. Additionally, imaging techniques such as chest CT scans are used to assess the severity of lung involvement. Treatment primarily involves supportive therapies, including oxygen therapy and mechanical ventilation.[8] Current treatment guidelines recommend the administration of anticoagulants, anti-inflammatory drugs, antiviral medications, broad-spectrum antibiotics, and anti-fungal agents for patients suspected or confirmed to have COVID-19.[9] However, despite implementing similar clinical interventions, patients may exhibit significant differences in recovery periods, clinical outcomes, or mortality rates.[10] The molecular mechanisms underlying these different outcomes remain largely unknown.[11]

Upon entering host cells, SARS-CoV-2 exploits the angiotensin-converting enzyme 2 (ACE2) receptor, binding to it through the receptor-binding domain of the spike (S) protein.[12] This interaction triggers fusion between the viral envelope and the host cell membrane via a Toll-like receptor 7 (TLR-7)-dependent pathway, allowing the release of the viral genome into the cytoplasm.[13] The virus uses its own RNA-dependent RNA polymerase to replicate its genome.[14] The replication-transcription complexes are formed within double-membrane vesicles, mediated by two large polyproteins (such as pp1a and pp1b), which encode non-structural proteins.[15] This process generates multiple subgenomic RNAs that encode structural proteins and accessory proteins.[16] Viral assembly and budding occur within smooth-walled vesicles of the endoplasmic reticulum-golgi intermediate compartment. Finally, the vesicles containing viral particles fuse with the cellular membrane, leading to the release of the virus through exocytosis.[17]

Numerous studies have established a strong correlation between innate immune response, adaptive immune development, and the severity of COVID-19.[18] Excessive inflammatory responses have been observed in various samples obtained from patients with COVID-19, including blood, nasopharynx, and bronchoalveolar lavage fluid. Elevated levels of cytokines such as IL6, tumor necrosis factor-alpha (TNF-α), and monocyte chemoattractant protein-1 have been detected, which can contribute to the development of SARS, coagulation disorders, and multi-organ failure.[4,19] IL6 is closely associated with respiratory failure and acute kidney injury, with its levels positively correlated with disease severity.[20] IL8 is linked to 30-day mortality.[21] IL-10, as an anti-inflammatory cytokine, plays a pivotal role in modulating immune responses. While IL-10 typically protects tissues by dampening excessive immune reactions, its abnormal elevation in COVID-19 patients may contribute to disease complexity and progression. Specifically, elevated IL-10 levels might impair the body's ability to clear the virus by suppressing the function of critical immune effector cells, while also potentially sustaining a chronic inflammatory state.

Although several studies have established prognostic models for patients with COVID-19 based on transcriptomic levels,[22–24] the overall effectiveness of these models has been found to be unsatisfactory. Although transcriptome sequencing data analysis can capture subtle biological changes that may be difficult to detect through clinical features, this type of data presents challenges due to its high dimensionality and complexity. It is particularly susceptible to various factors, including batch effects, which complicate data standardization. Additionally, obtaining gene expression data is generally more expensive and requires more time for experimentation and analysis compared to collecting clinical features. In contrast, clinical rapid testing is typically easier to perform and often exhibits high accuracy. As a result, models based on clinical features are usually more straightforward to understand and interpret, making them more accessible for clinicians in practical applications.

In this study, we conducted a comprehensive analysis of data from 189 patients with COVID-19 hospitalized in real-world settings, examining 149 clinical features to identify those most closely associated with clinical prognosis. Using these features, we developed a robust prognostic model with strong predictive capabilities. Our analysis identified IL-10 as a key factor influencing the prognosis of COVID-19 patients. Importantly, we validated these findings across multiple bulk transcriptomic and single-cell sequencing datasets, underscoring the robustness and reliability of our results. The identification of IL-10 as a significant prognostic biomarker has important clinical implications. Elevated IL-10 levels could serve as an early indicator of high-risk patients, allowing for timely intervention and more personalized treatment

strategies. This biomarker may also guide clinicians in adjusting therapeutic approaches, such as intensifying immuno-modulatory treatments for patients identified as high-risk. By providing a novel insight into the immune pathogenesis of COVID-19, our findings could inform clinical decision-making and contribute to improved patient outcomes.

Materials and Methods
Collection and Cleaning of Clinical Features
We retrieved data from the electronic medical record database of the Second Hospital of Hebei Medical University, consisting of 189 hospitalized patients admitted between November 1, 2022, and March 31, 2023. All patients were diagnosed with novel coronavirus pneumonia according to the "Diagnosis and Treatment Plan for Novel Coronavirus Pneumonia (Trial Tenth Edition)." The collected patient data included gender, age, smoking history, body mass index, symptoms at admission, vital signs, underlying diseases, blood gas analysis, laboratory examinations, presence of pathogenic bacteria, imaging examinations, electrocardiograms, presence of lower limb deep venous thrombosis, oxygen administration method, treatment, length of hospital stay, and final outcome. Regarding cytokine protein expression levels, we used the flow fluorescence technique to detect them. For the imaging examinations, we followed established protocols and divided each lung into three regions: the upper region (above the clavicle), the middle region (from the clavicle to the lower pulmonary vein), and the lower region (below the lower pulmonary vein). Each lung region was assigned a quantitative score based on the extent of involvement: 0 indicated no involvement; 1 indicated involvement of <25% of the area; 2 indicated involvement of 25–50% of the area; 3 indicated involvement of 50–75% of the area; and 4 indicated involvement of >75% of the area. Each patient had six lung regions, resulting in a maximum score of 24 (4 x 6 regions).[25] We performed data cleaning procedures to ensure data quality. First, we removed clinical features with >30% missing values, resulting in a final set of 149 valid indicators. For handling missing values, we used the mice package in R to perform imputation. Multivariate linear regression-based imputation was applied for numerical indicators, while multiple imputations by chained equations were used for categorical indicators.

Prognostic Model Building Based on Machine Learning Integration Framework
In this study, we aimed to develop a prognostic model using a machine learning integration framework. Specifically, we incorporated seven different machine learning algorithms, namely random survival forest (RSF), elastic network (Enet), least absolute shrinkage and selection operator (Lasso), Ridge, stepwise Cox, CoxBoost, and partial least squares regression for Cox (plsRcox). To evaluate their performance, we employed a ten-fold cross-validation strategy. To ensure model consistency and avoid over-reliance on a single feature, we required at least two features for each model, as models with only one feature may perform well on the training cohort but have limited generalization ability to new, unseen data. Through our integration process, we generated a total of 77 different combinations of machine learning algorithms. The performance of each model was assessed using Harrell's concordance index (C-index), a widely used measure of predictive accuracy for survival models. The model with the highest average C-index value was identified as having the best signature and selected for further analysis.

Gathering and Processing of Bulk Transcriptomic Data
For this study, we utilized bulk transcriptomic data obtained from the Gene Expression Omnibus (GEO) database, accessible at https://www.ncbi.nlm.nih.gov/geo/. Specifically, we analyzed two login identification numbers: GSE217948 and GSE179850. The GSE217948 dataset consisted of transcriptomic data from peripheral blood samples collected from 71 healthy individuals and 203 patients with COVID-19. Similarly, the GSE179850 dataset included transcriptomic data from 16 healthy individuals and 31 patients with COVID-19. Each dataset was analyzed individually to extract meaningful insights. To ensure the accuracy and reliability of our analysis, we employed specific methods in an R environment. Firstly, we used the "avereps" function available in the "limma" package to calculate average gene expression levels for replicate genes. Furthermore, we employed the "normalizeBetweenArrays" function to correct for any potential variability that may exist between the datasets.

Enrichment Analysis

The enrichment analysis of differential genes was conducted using the "ClusterProfiler" and "org.Hs.eg.db" packages. The EnrichGO and EnrichKEGG functions were used for enrichment analysis. Pathways with a P <0.05 were considered to be significantly enriched. Visualization of the results was accomplished using the "ggplot2" package.

Collection and Processing of Single-Cell Transcriptome Data

The single-cell transcriptome data used in this study were obtained from the GEO dataset, specifically the GSE217948 and GSE158055 datasets. For the GSE217948 dataset, quality control measures were performed within the R environment (version 4.1.2) using standard single-cell processing procedures. To process the count matrix, we used the Read10X function from the Seurat package (version 4.0.4) and converted it to the dgCMatrix format. To integrate all individual objects into an aggregate object, we employed the merge function. Additionally, the RenameCells function ensured unique cell labels. To filter out low-quality cells, we applied specific criteria. Genes expressed in <3 cells were removed, and cells expressing <200 genes were deleted. To equalize the total gene expression in each cell, we performed global-scaling normalization using the "LogNormalize" method with a scale factor of 10,000. For downstream analysis, we focused on the top 2000 variably expressed genes, which were identified using the FindVariableFeatures function. Unwanted sources of variation, such as unique molecular identifiers and percent mitochondrial content, were regressed using the ScaleData function with the "vars. to.regress" option. Principal component (PC) analysis was utilized to reduce the dimensionality of the dataset by incorporating highly variable features. The first 30 PCs were selected for further analysis. The Harmony method[26] was used to address the batch effects between samples. To visualize the cells in a lower-dimensional space, we applied the uniform manifold approximation and projection method, which effectively reduces the dimensionality of the data while preserving its local structure. Clustering analysis was performed based on edge weights between cells. We constructed a shared nearest-neighbor graph using the Louvain algorithm, implemented in the FindNeighbors and FindClusters functions. The resolution parameter in the FindClusters function was systematically adjusted between 0.1 and 1 to find the optimal clustering resolutions. We used the clustree function to evaluate the clustering trees at different resolutions and ultimately selected a resolution of 0.5, which provided clear and meaningful clustering results. To identify differentially expressed markers associated with the resulting clusters, we used the FindAllMarkers function. This function utilizes a default nonparametric Wilcoxon rank sum test with Bonferroni correction. To annotate all cells, we used cell surface markers and known genes from relevant literature, as well as the CellMarker database[27] (available at http://xteam.xbio.top/CellMarker/). For the GSE158055 dataset, we performed a rapid analysis using the online web version available at http://covid19.cancer-pku.cn/#/differential-expression.

Cellular Communication Network

For a comprehensive investigation of cell-cell interactions, we employed the "CellChat" (v1.0.0) R package.[28] This package provides us with a curated database of ligands, receptors, cofactors, and their respective interactions, which is publicly available at http://www.cellchat.org/. To analyze cell-cell communication, we projected the expressed ligand and receptor genes of each cell onto a reference communication network. By utilizing gene expression data, we inferred the probability of communication in each pathway. The CellChat package proved to be a versatile and user-friendly tool, enabling us to easily infer, analyze, and visualize cell-cell communication within our single-cell RNA sequencing data. For visualization purposes, we utilized the default parameters of the netVisual_bubble function, which allowed us to visually represent the communication pathways with bubble plots.

Statistical Analysis

To perform our statistical analyses, we used R (version 4.1.2). We considered a *P*-value of <0.05 to be statistically significant.

Results

Prognostic Model Building Based on Machine Learning Integration Framework

A total of 189 COVID-19 patients were included in the study. Patient clinical and demographic characteristics are shown in Table 1. Our workflow, shown in Figure 1, began with conducting univariable Cox regression analysis on all 149 features to identify prognostic indicators. Among these features, 43 were found to be prognostically relevant ($P < 0.05$), comprising 13 protective features (hazard ratio [HR] <1) and 30 hazard features (HR >1) (Table S1). To construct an optimal prognostic model, we integrated these 43 features into our machine-learning process. For validation purposes, we randomly divided the 189 case samples into two cohorts, namely the training and validation cohorts, at a ratio of 1:1. The comparative 43 features of patients in training cohorts, validation cohorts and all patients are described in Table S2. The model training was exclusively conducted in the training cohort, and we subsequently evaluated the accuracy and validity in both the validation and entire cohorts. Using ten-fold cross-validation, we used a combination of seven machine learning algorithms, namely RSF, Enet, Lasso, Ridge, stepwise Cox, CoxBoost, and plsRcox. One algorithm was used for feature filtering, while another was utilized for constructing a prognostic model. Through this approach, we successfully generated 77 different model combinations within the training cohort. For each model, we calculated the C-index across all validation datasets, selecting the model with the highest average C-index as the optimal choice. Among all the model combinations, the combination of Lasso and RSF demonstrated superior prognostic prediction performance (Figure 2A). Consequently, we named this model the COVID-19-related prognostic model (CRM). The C-index values of the CRM

Table 1 Clinical Characteristics of Study Participants

Variable	Total COVID-19 (n=189)
Age years,median [IQR]	72(61–82)
Female (%)	128(67.7)
Smoking (%)	63(33.3)
Mechanical ventilation(%)	15.3 (29)
Hospital LoS, median [IQR]	17 (12–24)
Comorbidities, n (%)	
Hypertension	88(46.6)
Coronary Heart Disease	60 (31.7)
Diabetes Mellitus	58 (30.7)
Cancer	23 (12.2)
BMI, median [IQR]	24.2 (20.8–26.4)
Symptoms, n (%)	
Fever	123(65)
Cough	153(80.9)
Fatigue	47(24.8)
Dyspnoea	82(43.3)
Routine clinical data at inclusion, median [IQR]	
Leukocytes ×10^9/L	6.7 (5–9.2)
Lymphocytes ×10^9/L	0.72 (0.5–1.25)
Neutrophils×10^9/L	79.5(67.2–88.6)
Monocytes ×10^9/L	0.4 (0.29–0.6)
Platelets ×10^9/L	201 (155–281)
Procalcitonin µg/L	0.05 (0.03–0.18)
CRP, mg/L	32.7 (8.46–101)
Fibrinogen g/L	4.28 (3.64–4.88)
D-dimer mg/L	0.46 (0.2–0.99)
PaO2, mmHg	76.8 (61.5–94.6)
PaCO2, mmHg	33.7 (29.4–37.8)

Note: Clinical biochemical and haematological routine laboratory assays.
Abbreviations: IQR, interquartile range; LoS, Length of stay; BMI, body mass index; CRP, C-reactive protein.
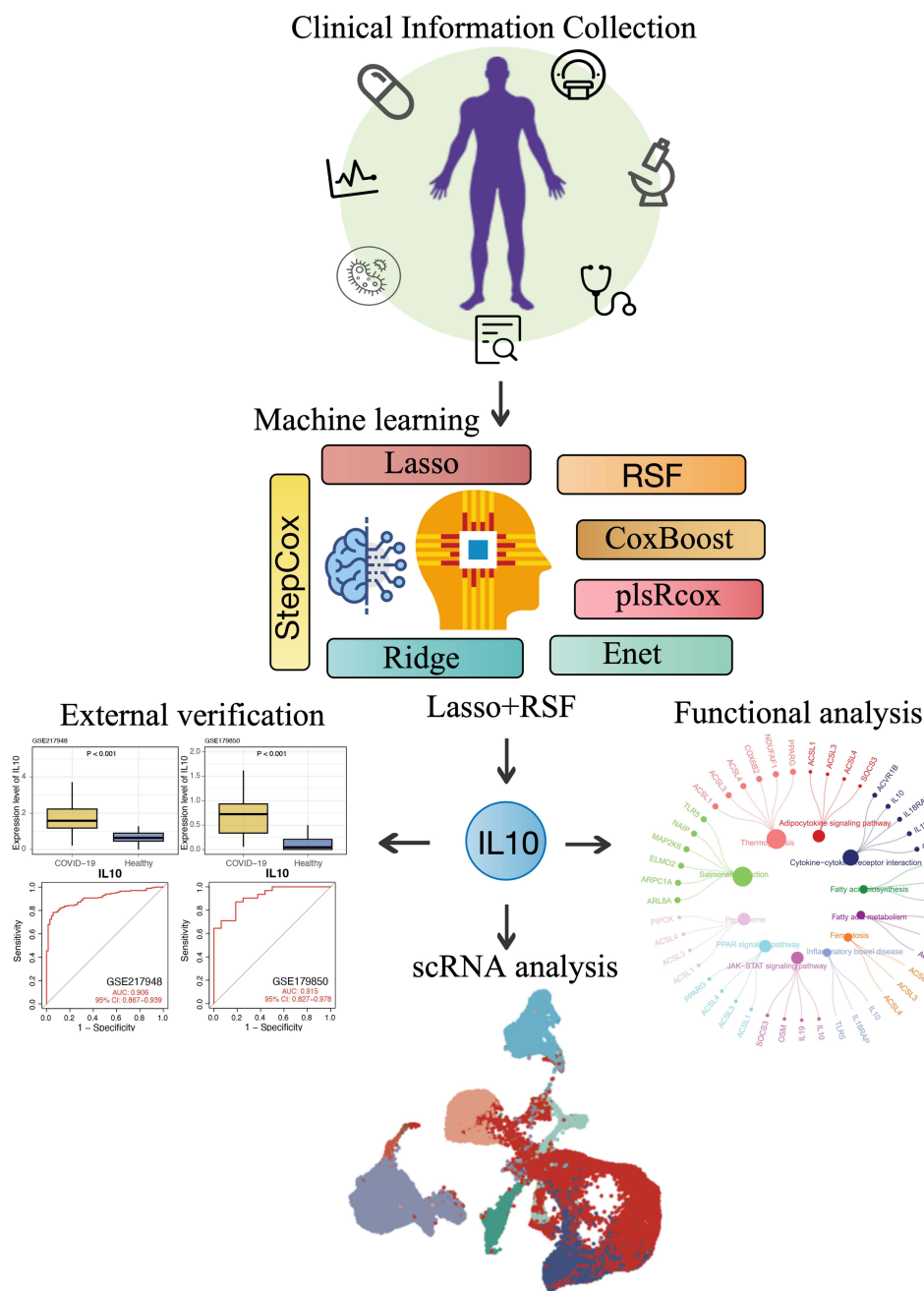
Figure I Overview of all analytical processes conducted in this study.

were 0.918, 0.911, and 0.914 in the training cohort, validation cohort, and entire cohort, respectively, indicating the robustness and reliability of the CRM. Using the predict function within the CRM, we calculated a risk score for each patient. The Lasso regression identified four features for inclusion in the CRM (Figure 2B and C). Subsequently, RSF assessed the importance of these four features, ranking them in the following order of importance: IL-10, alpha-hydroxybutyric dehydrogenase, pro-brain natriuretic peptide, and unconsciousness (Figure 2D and E).

Accuracy of CRM in Predicting the Prognosis of Patients with COVID-19

In order to stratify the patients into high-risk and low-risk groups, we used the "srvminer" package to determine the optimal threshold value based on individual risk scores in the training cohort. Our analysis revealed that the high-risk patients had significantly reduced overall survival (OS) compared to the low-risk patients (Figure 3A). This finding was
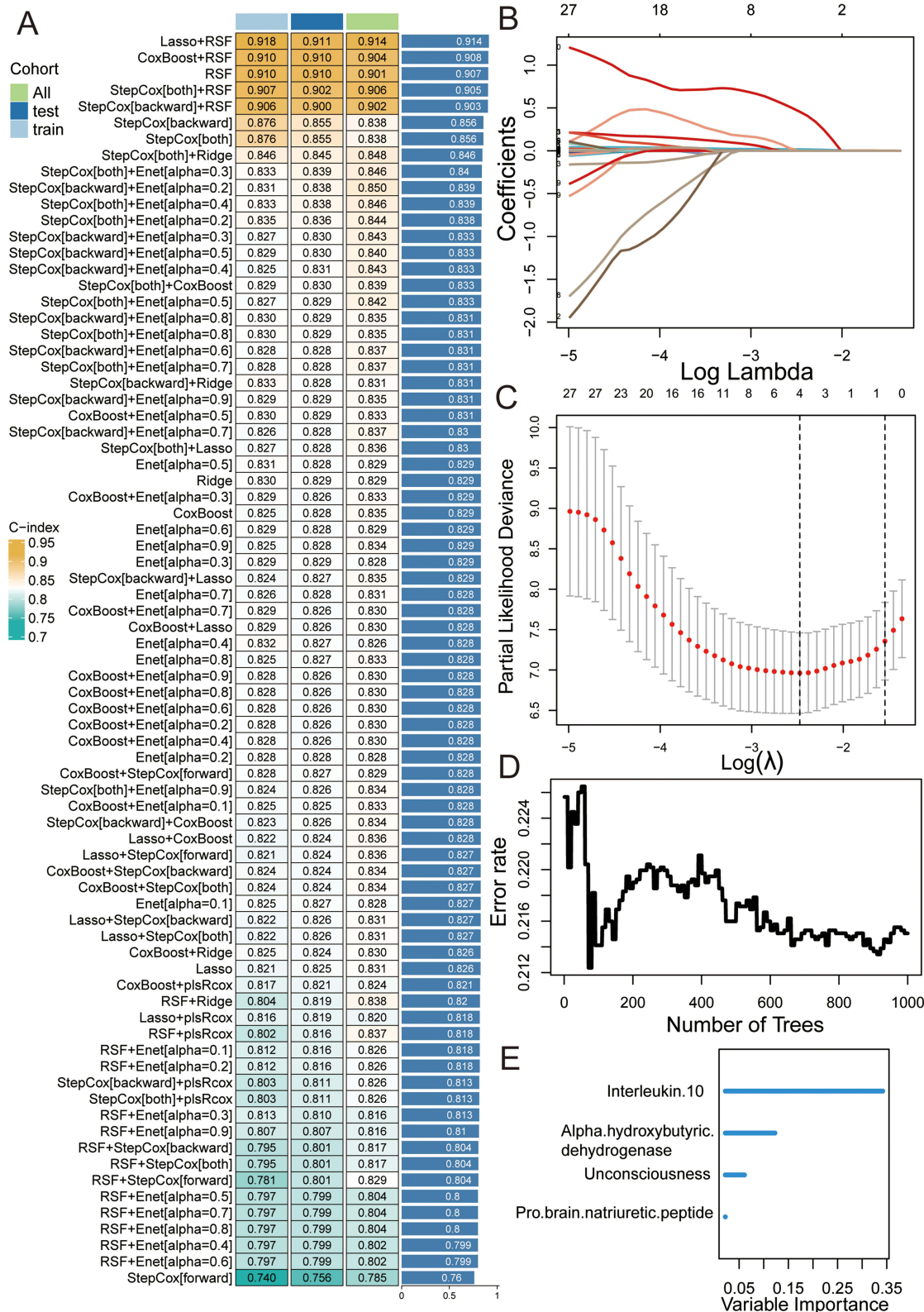
Figure 2 Establishment of the machine learning-integrated prognostic model. (**A**) C-index values of 77 machine learning combinations represented by a color spectrum from green to yellow, indicating increasing C-index values. Each row represents a signature, and each column represents a cohort. (**B** and **C**) Lasso regression used for signature selection analysis. (**D**) RSF analysis to select minimum number of trees. (**E**) Assessment of signature importance using the RSF algorithm.
Abbreviations: RSF, random survival forest; Lasso, least absolute shrinkage and selection operator; C-index, concordance index.
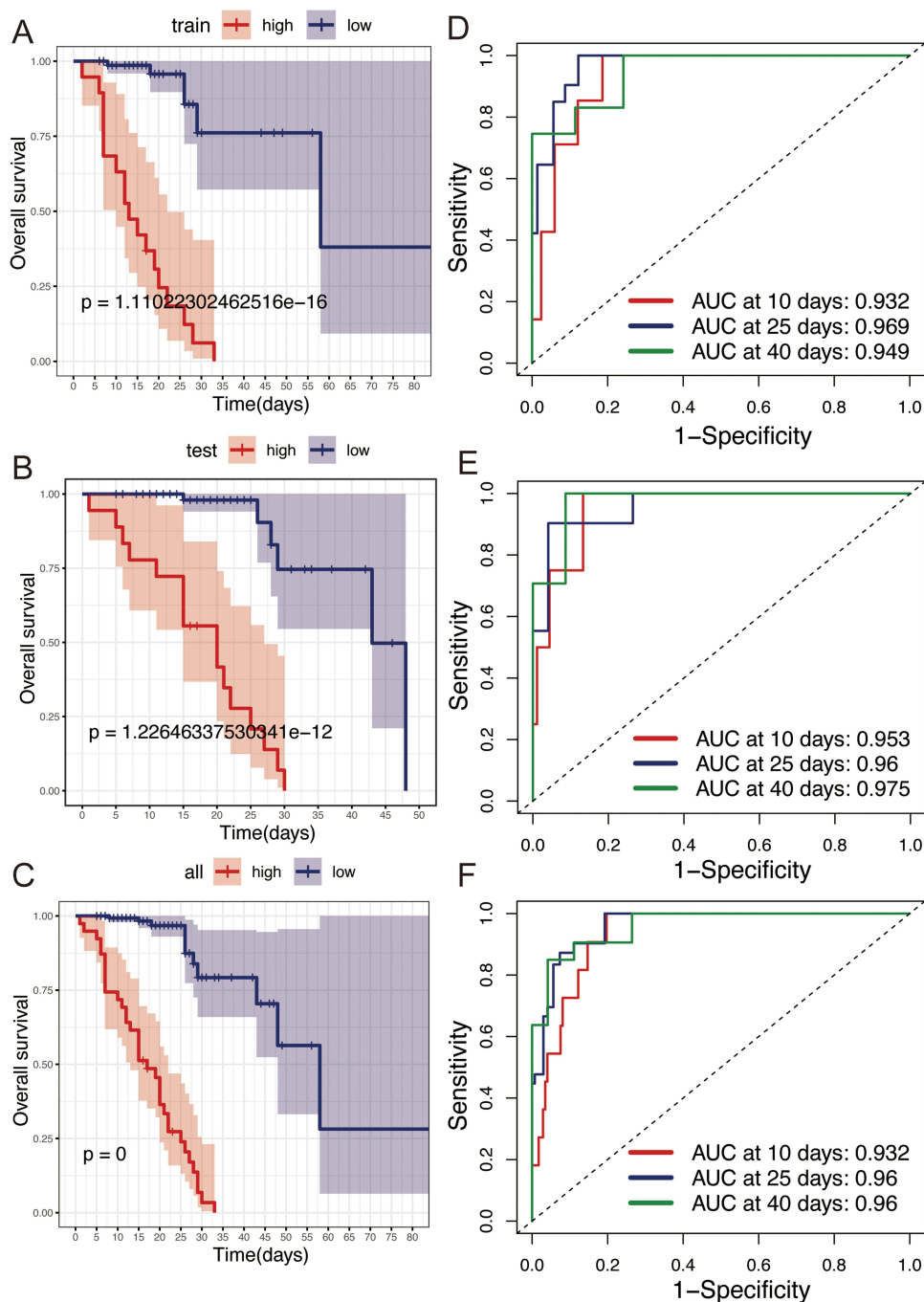
Figure 3 Evaluation of the effectiveness and accuracy of the CRM. (**A**) Survival analysis comparing high- and low-risk patients in the training cohort. (**B**) Survival analysis comparing high- and low-risk patients in the validation cohort. (**C**) Survival analysis comparing high- and low-risk patients in the entire cohort. (**D**) ROC curve analysis of risk scores for the training cohort. (**E**) ROC curve analysis of risk scores for the validation cohort. (**F**) ROC curve analysis of risk scores for the entire cohort.
Abbreviations: CRM, coronavirus disease 2019-related prognostic model; ROC, receiver operating characteristic.

consistent and replicated in both the validation cohort (Figure 3B) and the entire cohort (Figure 3C), providing strong evidence for the association between high-risk scores and poorer prognostic outcomes. To further assess the predictive performance of the risk scores, We performed a receiver operating characteristic (ROC) analysis, and the results demonstrated the excellent discriminative ability of the risk scores at different time points. In the training cohort, the area under the curve (AUC) values for risk scores at 10, 25, and 40 days were 0.932, 0.969, and 0.949, respectively (Figure 3D). Similarly, the AUC values for risk scores at the same time points were 0.953, 0.960, and 0.975 in the

validation cohort (Figure 3E) and 0.932, 0.960, and 0.960 in the entire cohort (Figure 3F), respectively. These results demonstrate the robust effectiveness and accuracy of the CRM in predicting patient outcomes.

Stable Diagnostic and Prognostic Value of IL-10

The CRM identified the expression of IL-10 as an important prognostic factor for patients. IL-10, a cytokine regulator, plays a role in modulating immune responses by exerting anti-inflammatory and immunosuppressive effects. In the case of patients with COVID-19, high levels of IL-10 can hinder the host immune response, preventing viral clearance and worsening the condition of the patient. Based on these findings, we conducted a comprehensive analysis of IL-10. First, our analysis revealed that the expression of IL-10 was consistently higher in the high-risk group compared to the low-risk group within training (Figure 4A) and validation cohorts (Figure 4B), as well as across the entire cohort (Figure 4C). Moreover, patients with elevated expression of IL-10 in these cohorts exhibited significantly reduced OS compared to patients those with lower expression of IL-10 (Figures 4D–F). To further assess the diagnostic potential of IL-10 in distinguishing healthy individuals from patients with COVID-19, we examined the GSE217948 and GSE179850 transcriptomic datasets. GSE217948 consisted of transcriptomic data from 71 healthy samples and 203 patients with COVID-19, while GSE179850 included transcriptomic data from 16 healthy individuals and 31 patients with COVID-19. Our analysis of both datasets revealed a significant elevation in the expression of IL-10 in patients with COVID-19 compared to healthy individuals (Figure 4G). Furthermore, the ROC curve showed strong diagnostic performance for IL-10, with an AUC of 0.906 in the GSE217948 dataset and 0.915 in the GSE179850 dataset (Figure 4H). In summary, our findings highlight the critical role of elevated levels of IL-10 in patients with COVID-19, indicating its association with a poor prognosis.

IL-10-Related Functional Analysis

To further explore the functional pathways potentially influenced by IL-10, we conducted a correlation analysis between IL-10 and the expression levels of all other genes in patients with COVID-19 using the GSE217948 dataset. Our analysis revealed a positive correlation between the expression of IL-10 and several genes associated with the inflammatory response and immune regulation (Figure 5A). For instance, *MS4A4A*, which is involved in signal transduction and cellular vitality, could potentially influence immune cell activity. IL-19 and SOCS3, both anti-inflammatory cytokines, may work in conjunction with IL-10 to modulate immune responses. TLR5, responsible for pathogen recognition, exerts immunomodulatory effects. COX6B2, a protein related to oxidative phosphorylation, is involved in cellular energy metabolism. Additionally, CD163, a receptor expressed on macrophages and monocytes, suggests that IL-10 may impact these cells, potentially promoting the formation of M2 macrophages. To gain deeper insights into the functional impact of IL-10, we performed gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses using the top 100 co-expressed genes with IL-10. GO enrichment analysis (Figure 5B) revealed enrichment in various metabolic and developmental processes. Metabolic pathways such as the ribonucleotide metabolic process, the ribose phosphate metabolic process, and the long-chain fatty-acyl-CoA metabolic process suggest that IL-10 may regulate immune responses and antiviral capabilities through the regulation of metabolic pathways. Moreover, pathways related to the development of the placenta and the morphogenesis of the embryonic placenta shed light on the potential role of IL-10 in governing cell proliferation and differentiation processes. In terms of cellular components, the significant enrichment of pathways like the mitochondrial outer membrane, peroxisome, trans-golgi network, and nucleolar exosome further supports the notion that IL-10 may be involved in cellular defense against viruses, as well as cell proliferation and differentiation processes, by regulating and repairing cellular organelles. In the molecular function category, enriched activities such as carbohydrate kinase activity, fatty acid ligase activity, acid-thiol ligase activity, cysteine-type endo-peptidase inhibitor activity, and calmodulin-dependent protein kinase activity suggest that IL-10 could participate in metabolic pathways and cell proliferation processes by modulating enzyme activity. KEGG enrichment analysis (Figure 5C) highlights pathways such as thermogenesis, peroxisome proliferator-activated receptor (PPAR) signaling, fatty acid metabolism, and inflammatory bowel disease, all of which are relevant to metabolic disorders. This finding aligns with the results of the GO analysis. Additionally, enriched pathways such as *Salmonella* infection and ferroptosis provide insights into the potential role of IL-10 in modulating immune responses. Collectively, our findings suggest that
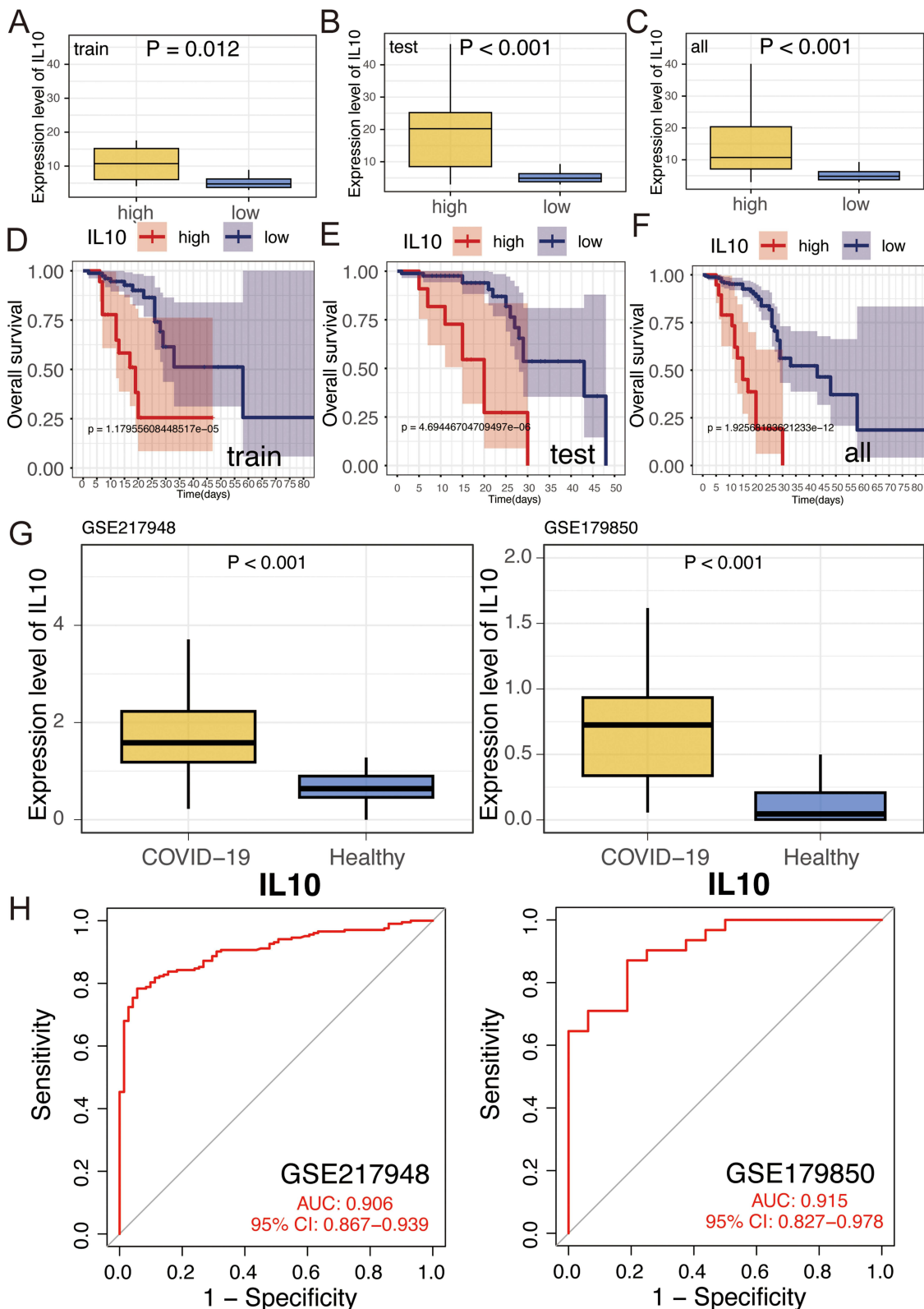
Figure 4 Analysis of the diagnostic and prognostic value of IL-10. (**A–C**) Expression of IL-10 in high- and low-risk patients in the training, validation, and entire cohorts. (**D–F**) Survival analysis comparing groups expressing high and low levels of IL-10 in the training, validation, and entire cohorts. (**G**) Box plots depicting differential expression of IL-10 between healthy individuals and patients with COVID-19 in the GSE217948 and GSE179850 datasets. (**H**) ROC curves illustrating the differentiation of IL-10 between healthy individuals and patients with COVID-19 in the GSE217948 and GSE179850 datasets.

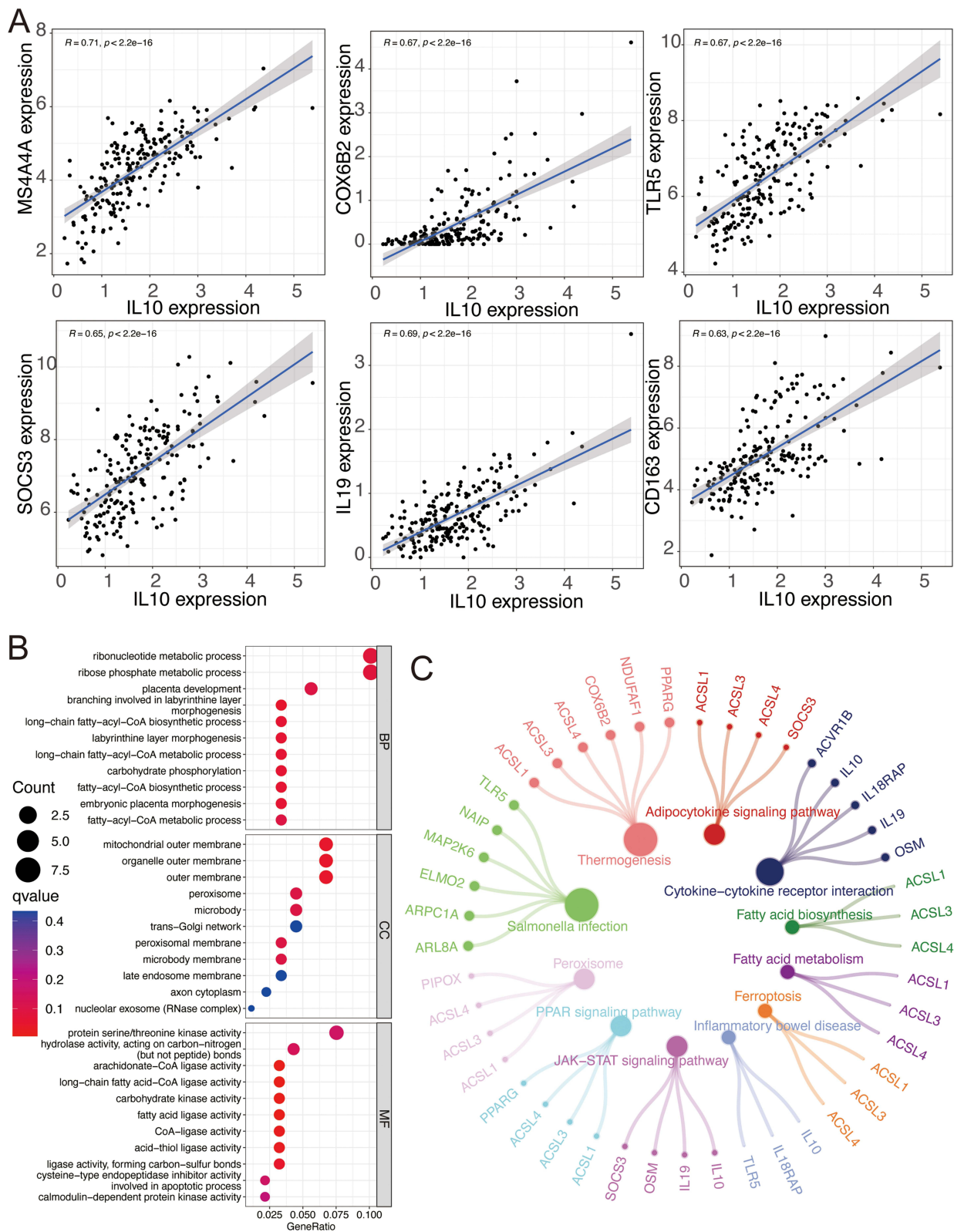**Abbreviations**: IL, interleukin; COVID-19, coronavirus disease-2019; ROC, receiver operating characteristic.

Figure 5 Functional analysis related to IL-10. (**A**) Pearson correlation analysis of genes co-expressed with IL-10. (**B**) GO enrichment analysis. (**C**) KEGG enrichment analysis.
Abbreviations: IL, interleukin; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

IL-10 may influence the prognosis of patients with COVID-19 by affecting pathways related to metabolism, proliferation and differentiation of cells, as well as the regulation and repair of cellular organelles and immune responses. These results establish a foundation for future studies exploring the immunopathological mechanisms of COVID-19.

Single-Cell Sequencing Analysis: Monocytes as the Primary Source of the Elevated IL-10 in Patients with COVID-19

Although our clinical data and transcriptomic analysis have demonstrated the robust diagnostic and prognostic value of IL-10, the specific source of IL-10 in patients with COVID-19 has remained elusive. Bulk transcriptomic analysis tends to mask crucial differences among cellular subpopulations. However, the emergence of single-cell sequencing technology presents an opportunity to uncover the intrinsic features of patients with COVID-19 at the single-cell level. In this study, we obtained single-cell transcriptome data from peripheral blood samples of 11 patients with COVID-19 and four healthy individuals, sourced from the GSE149689 dataset. Following quality control measures, we included a total of 70,559 cells for subsequent analyses. Using cell surface markers, we annotated all cells (Figure 6A) and identified nine distinct cell types (Figure 6B). These cell types encompassed B cells (CD79A and MS4A1), dendritic cells (DC) (CLEC10A, CST3, HLA-DPA1, and HLA-DQA1), plasmacytoid DC (pDC) (ALOX5AP, APP, CCDC50, IRF7, IRF8, and LILRA4), erythroid cells (HBA1, HBA2, HBB, and AHSP), innate lymphoid cells (ILC) (GNLY, GZMA, GZMM, and NKG7), megakaryocytes (PPBP, GNG11, MYL9, and PF4), monocytes (S100A9, S100A11, AIF1, and S100A8), plasma cells (IGHA1, IGHG1, IGHG3, IGHG4, and IGKC), and T cells (CD3D, CD3E, and IL7R). Consistently, our analysis revealed significantly elevated expression levels of IL-10 in patients with COVID-19 compared to healthy samples (Figure 6C), aligning with the findings from bulk transcriptomic data analysis. Notably, we observed that monocytes exhibited the highest expression of IL-10 compared to other cell types (Figure 6D), suggesting that monocytes serve as the primary source of elevation of IL-10 in patients with COVID-10. To validate our findings externally, we further analyzed the GSE158055 dataset, which included 284 single-cell transcriptomic profiles collected from 196 individuals, including 22 patients with mild or moderate symptoms, 54 hospitalized patients with severe symptoms, and 95 recovered convalescent individuals, as well as 25 healthy individuals. The dataset encompassed samples from various tissue types, such as peripheral blood mononuclear cells, bronchoalveolar lavage fluid, and pleural effusion sputum. Similarly, our analysis demonstrated that the expression of IL-10 was highest at the monocyte level (Figure 6E). Additionally, expression levels of IL-10 were elevated in nucleic acid test-positive specimens compared to negative ones (Figure 6F). Furthermore, the healthy individuals exhibited the lowest expression level of IL-10, while the progression group showed the highest expression, with levels of IL-10 partially recovering in the convalescent group (Figure 6G). The levels of IL-10 gradually increased from the normal level to mild, moderate, severe, or critical levels (Figure 6H). We also discovered a gradual increase in the expression levels of IL-10 with age (Figure 6I), which may contribute to the relatively severe condition of COVID-19 in elderly individuals. In conclusion, our single-cell analysis findings indicate that monocytes are the primary source of increased IL-10 in patients with COVID-19.

Monocytes Expressing High Levels of IL-10 are Involved in Immune Cell Communication in Patients with COVID-19

Cellular communication plays a crucial role in understanding the complexity of the immune microenvironment in patients with COVID-19, as changes in this microenvironment can influence disease progression. However, bulk transcriptomic data fails to capture this aspect adequately. To address this, we divided all monocytes into two groups based on the expression of IL-10: high- and low-expressing monocytes (Figures 7A–C). Subsequently, we examined changes in cellular communication between these two types of cells and other cells (Figures 7D and E). Our analysis revealed several specific changes in signaling between monocytes expressing high levels of IL-10 and other cells compared to those monocytes expressing low levels of IL-10. Notably, the activation of ANXA1-FPR2 signaling between monocytes expressing high levels of IL-10, DCs, ILCs, pDCs, T cells, and other immune cells suggests a critical role in modulating the immune response, as this pathway is associated with the resolution of inflammation. The activation of MIF-(CD74+CXCR4) signaling between B and T cells as ligand cells and monocytes expressing high
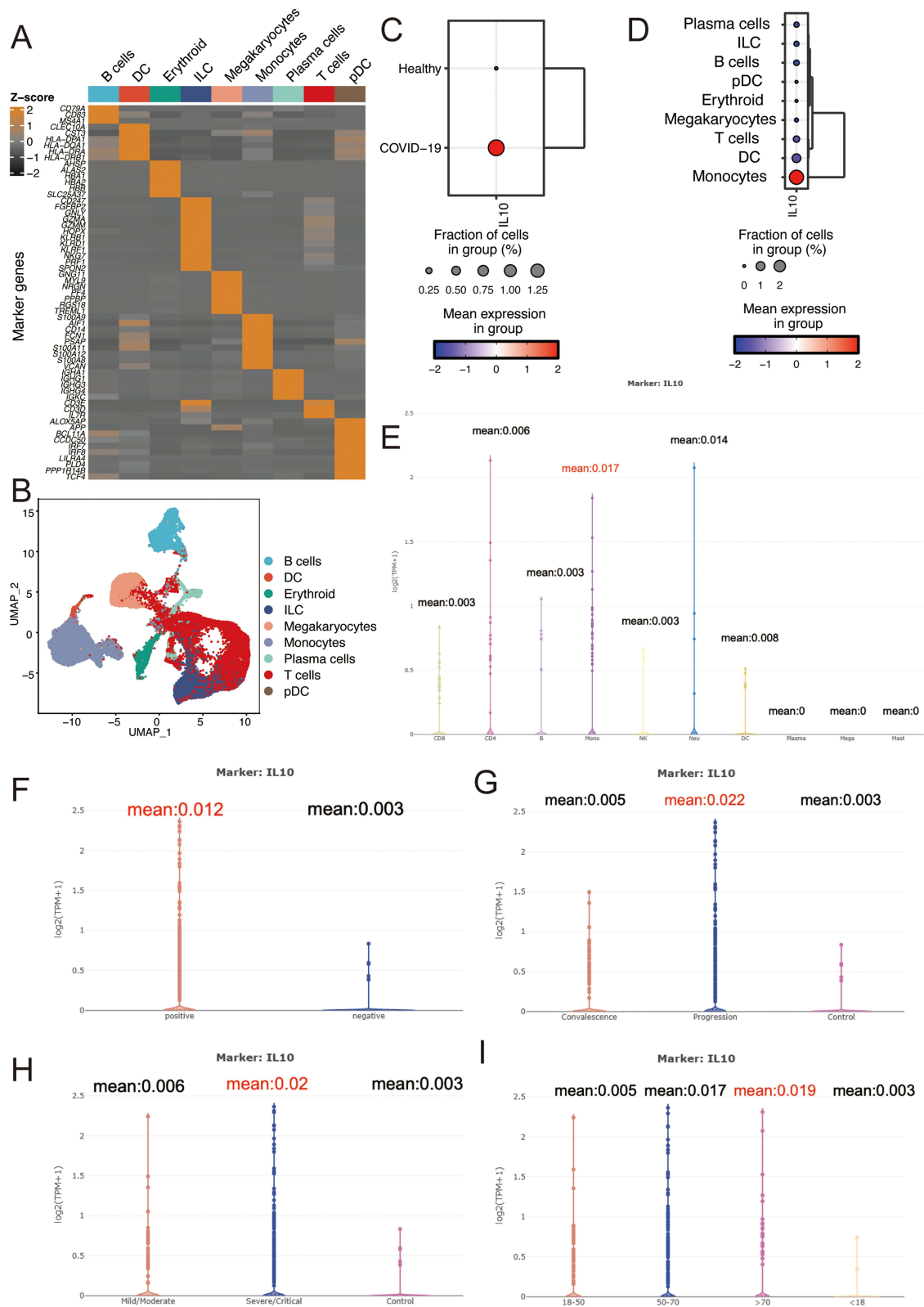
Figure 6 Differential expression analysis of IL-10 at the single-cell level. (**A**) A heatmap showing marker gene expression used for cell annotation in each cell in the GSE149689 dataset. (**B**) Dimensional landscape showing cell annotation in the GSE149689 dataset. (**C**) Differential expression of IL-10 in patients with COVID-19 and healthy individuals within the GSE149689 dataset. (**D**) Differential expression of IL-10 in different cell types in the GSE149689 dataset. (**E**) Differential expression of IL-10 in different cell types in the GSE158055 dataset. (**F–I**) Differential expression of IL-10 in different clinical feature subgroups within the GSE158055 dataset.
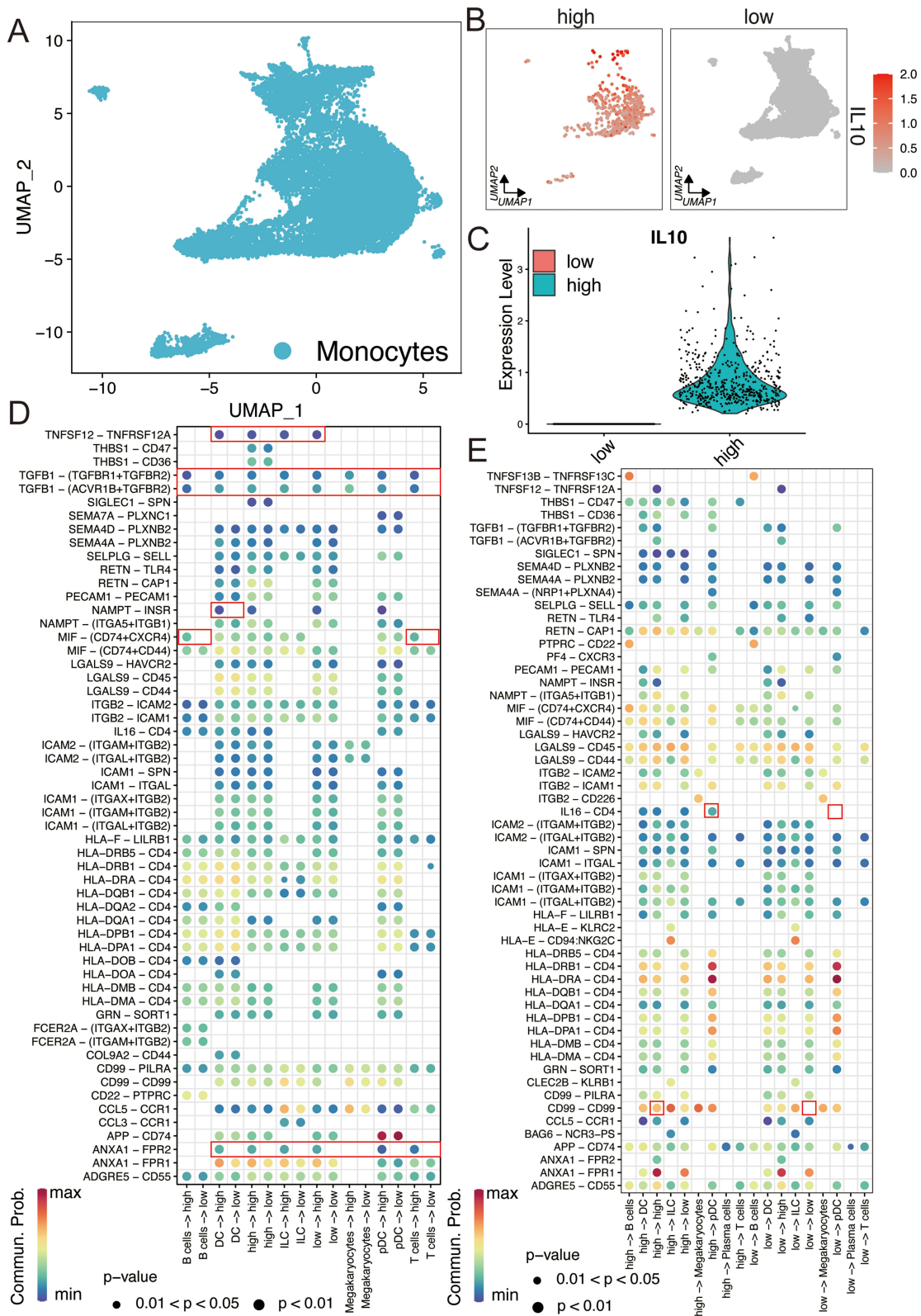**Abbreviation**: COVID-19, coronavirus disease-2019.

Figure 7 Analysis of cell communication. (**A**) The re-dimensionalization landscape of all monocytes after re-extraction. (**B** and **C**) Splitting of single monocytic cells into groups with high or low expression of IL-10 based on the expression of IL-10; (**D** and **E**) Communication analysis between the two groups expressing IL-10 using CellChat software, with "high" representing single monocytic cells with high expression of IL-10, and "low" representing single monocytic cells with low expression of IL-10.
Abbreviation: IL, interleukin.

levels of IL-10 as receptor cells indicates a potential involvement in cellular communication. Similarly, the activation of NAMPT-INSR signaling between DCs as ligand cells and monocytes expressing high levels of IL-10 as receptor cells highlights another interaction. Additionally, the activation of transforming growth factor (TGF) B1-(TGFBR1 + TGFBR2) and TGFB1-(ACVR1B + TGFBR2) signaling between other cells and monocytes expressing high levels of IL-10 is consistent with previous studies suggesting a role for TGF-β signaling in promoting immunosuppression. The activation of tumor necrosis factor superfamily (TNFSF) 12-TNFRSF12A signaling between DCs, ILCs, and monocytes expressing high levels of IL-10 may contribute to the activation and differentiation of T cells. Finally, we observed enhanced autocrine signaling of CD99-CD99 in monocytes expressing high levels of IL-10 compared to those expressing low levels of IL-10, which may impair cellular immune function and hinder humoral immune responses, potentially increasing susceptibility to infection. CD99-CD99 signaling is involved in vital cellular processes such as cell adhesion, cell-mediated toxicity, and intercellular interactions. Consequently, enhanced autocrine signaling of CD99-CD99 in monocytes expressing high levels of IL-10 could be related to the disease deterioration and poor prognosis in patients with COVID-19. In summary, these findings shed new light on the intricate cellular communication among immune cells in patients with COVID-19 and altered immune microenvironments. Our study provides a comprehensive understanding of the specific changes in cellular communication associated with monocytes expressing high levels of IL-10 in patients with COVID-19. These results have the potential to guide the development of novel therapeutic approaches and aid in the identification of biomarkers for predicting disease progression.

Discussion

In this study, our aim was to gain a deeper understanding of the disease status of hospitalized patients with COVID-19 and identify novel biomarkers that could assist in predicting patient prognosis. Initially, we analyzed real-world datasets to examine various clinical features of patients with COVID-19, including imaging, drug treatment, and peripheral blood characteristics. Subsequently, we developed a robust clinical feature model using consensus frameworks in machine learning to predict patient prognosis effectively. Notably, our model identified IL-10 as a crucial clinical indicator that significantly influences patient prognosis, a finding consistently validated in both bulk transcriptome and single-cell sequencing datasets. Moreover, our analysis of single-cell transcriptome data revealed that monocytes in the peripheral blood of the patients were the primary source of IL-10. These monocytes expressing high levels of IL-10 were involved in cell-to-cell communication within the immune microenvironment of the patient. Collectively, our findings highlight the importance of IL-10 as an important biomarker, emphasizing the need to test for levels of IL-10 early in inpatient admission and throughout the entire hospitalization period.

Previous studies have highlighted the significant changes in levels of various cytokines experienced by patients with COVID-19.[29] The infection triggers the immune system, resulting in variations in immune cell numbers and proportions, leading to the secretion of various cytokines that promote and regulate immune responses.[30] However, the specific cytokines that play a crucial role remain unclear. It has been observed that COVID-19 elevates the levels of cytokines in the body, including IL-10, IL-6, TNF-α, interferon-gamma (IFN-γ), and IL-1β, among others.[31–34] These cytokines have the capacity to regulate the inflammatory response of the body, potentially leading to pulmonary and cardiovascular complications.[35] While some studies suggest that IL-10 is not a risk factor for COVID-19, they emphasize the association of higher levels of IL-6 with the severity of the disease and mortality rates.[36] Conversely, some studies indicate that the combined action of TNF-α and IFN-γ can trigger extensive apoptosis, pyroptosis, and necroptosis of immune cells in the bodies of patients with COVID-19.[37] Furthermore, studies demonstrate that compared to previous outbreaks such as SARS and Middle East respiratory syndrome (MERS), COVID-19 exhibits higher levels of IL-10, indicating that the immune response to COVID-19 exhibits higher levels of IL-10, suggesting potentially distinctive immunomodulatory characteristics.[8] Overall, COVID-19 infection is characterized by significant changes in the levels of cytokines, which have implications for the severity and prognosis of the disease.[1] Understanding these changes is crucial for the development of more effective treatment strategies and the prediction of the risk of infection following COVID-19.

In our analysis, we investigated the genes associated with the expression of IL-10 and identified the top six co-expressed genes: *MS4A4A, IL-19, SOCS3, TLR5, COX6B2*, and *CD163*. Notably, *MS4A4A* and IL-19 exhibited high co-expression coefficients with IL-10, reaching approximately 0.7. Previous studies have also confirmed the relevance of

these genes to COVID-19. MS4A4A is a transmembrane protein widely expressed in immune cells, particularly monocytes and B cells.[38] It plays a role in immune regulation during viral infections. A large-scale gene association study targeting COVID-19 revealed that *MS4A4A*, located on chromosome 11q12.2, is associated with severe COVID-19 infection.[39] Expression of *MS4A4A* is affected by COVID-19 in monocytes and B-cells, with further downregulation observed in patients with severe infection.[40] Additionally, *MS4A4A* has been found to regulate cellular apoptosis and activation of T cells, indicating its significance in the immune response to COVID-19.[41] Decreased expression of *MS4A4A* in patients with COVID-19 may lead to a compromised antiviral immune response, potentially exacerbating the disease.[42] Moreover, other studies have suggested that *MS4A4A* can regulate inflammatory responses, which are major contributors to pathological damage caused by lung inflammation in COVID-19, leading to death.[42] Therefore, *MS4A4A* may play a crucial role in suppressing pathological inflammatory responses, and its decreased expression may be associated with the severity of the disease.

Similarly, IL-19 is also regulated during severe COVID-19 infections and plays a vital role in regulating immune and inflammatory responses, proliferation, and differentiation.[43] During COVID-19 infection, immune cells produce excessive inflammatory cytokines, leading to an overactive immune response. IL-19 may act as a negative regulator, mitigating this excessive immune response.[44] For example, IL-19 can inhibit the recognition of pathogens by alveolar macrophages and reduce macrophage activity.[45] While this may increase the risk of bacterial infections in certain cases, it could help alleviate lung inflammation and atelectasis in COVID-19. Furthermore, studies suggest that IL-19 is involved in immune response and antiviral immunity.[46] IL-19 can promote viral clearance, facilitate pathogen-mediated cross-interference, and regulate the polarization state of immune cells, leading to reduced inflammatory responses.[47] Patients with severe COVID-19 infection often experience a cytokine storm, which is characterized by heightened inflammation and tissue damage due to an overactive immune response.[37] IL-19 may contribute to the occurrence and development of cytokine storms,[48] making it an important regulator of immune responses in COVID-19. By inhibiting excessive immune reactions and alleviating inflammatory responses, IL-19 plays a crucial role in mitigating the severity of COVID-19 infections. In summary, IL-19 holds significant potential in the context of severe COVID-19 infections. While our understanding of its precise mechanism is still evolving, studying the role of IL-19 in COVID-19 provides valuable insights for better prevention and treatment of COVID-19.

We investigated the impact of COVID-19 on cellular functions and signaling pathways related to IL-10. Our enrichment analysis revealed that IL-10 may influence various metabolic pathways and immune responses. These findings are consistent with previous studies that have explored the role of IL-10 function in similar contexts.[49] Specifically, in the case of pulmonary viral infections, IL-10 can inhibit the development of IL-17-producing cells that can cause tissue damage, suppress innate inflammatory responses, and potentially reduce the production of cytokines.[50] The increased levels of IL-10 observed in patients with COVID-19 may directly result from the progression of lung inflammation and serve as a part of feedback inhibition circuits that counteract inflammatory responses.[50] Furthermore, elevated levels of IL-10 can serve as blood-based biomarkers for patients with COVID-19 who experience worsening lung damage.[51] However, individuals with higher levels of IL-10 may exhibit adaptive immune deficiencies, which have been associated with long-term, persistent symptoms of viral infection.[52] This overproduction of IL-10 could compromise the immune response to recurrent bacterial infections, leading to secondary pneumonia and playing a significant role in the progression of COVID-19 in patients.[53]

In subsequent studies, we categorized monocytes from the single-cell data into two groups based on their expression levels of IL-10. The analysis revealed notable differences in cell communication between cells with high expression of IL-10 and those with low expression, particularly in the TNFSF and TGFβ signaling pathways. This finding is consistent with a previous study demonstrating the involvement of various TNFSF proteins, such as CD40L, FasL, gitr, and OX40, in COVID-19. CD40L, for example, has been associated with severe COVID-19 infections and is involved in B cell and multiple antibody responses.[54] FasL, expressed by CD8 T cells, may contribute to COVID-19 by inducing cell apoptosis.[55] Additionally, GITR and OX40, expressed in T cells and plasma cells, may also have implications for COVID-19.[56] Consequently, further extensive and in-depth research is needed to comprehensively understand the role of TNFSF in COVID-19, offering valuable insights for the treatment and prevention of COVID-19. TGFB1, another significant factor, participates in the regulation of the immune response in patients with COVID-19. Studies have

shown that in COVID-19, virus-induced immune cells release excessive inflammatory cytokines, leading to an extreme activation of the immune system, characterized by the activation and infiltration of inflammatory cells, which results in an excessive inflammatory response and lung damage.[57] TGFB1 can regulate the activation of inflammatory cells and suppress the inflammatory response, thus playing a protective role.[58] A study observed a decrease in the levels of TGFB1 in the serum of patients with severe COVID-19, possibly related to the excessive activation of immune cells caused by the viral infection.[59] Furthermore, TGFB1 also plays an important role in the onset and progression of COVID-19-related lung injury.

Studies have shown that the COVID-19 virus impacts the upregulation of *ACE2* and *TMPRSS2*, receptors associated with virus infection. This leads to the release of inflammatory mediators, which further activate the immune response, ultimately resulting in damage to lung parenchymal tissue.[57] Within this process, the regulation of the expression of TGFB1 influences the functionality of immune cells, including inflammatory cells, T cells, B cells, macrophages, and fibroblasts, thereby affecting their involvement in the onset and progression of lung injury.[60] Overall, TGFB1 plays a crucial regulatory role in COVID-19, particularly in severe cases, by regulating various aspects such as immune and inflammatory cells, lung injury, and prognosis. However, there is currently limited research on the specific mechanisms and therapeutic effects of TGFB1 in COVID-19. Therefore, further investigation is necessary to enhance our understanding of the association between this protein and severe COVID-19 infection, enabling better guidance for the treatment and prevention of COVID-19 in patients.

In conclusion, we established a prognostic model for patients with COVID-19 using a machine-learning consensus framework, incorporating real-world clinical features. Notably, our study identified IL-10 as a critical prognostic factor. These findings offer valuable insights into the intricate pathogenesis and progression of COVID-19. They have the potential to aid in the identification of high-risk patients with COVID-19 at fatal risk and may enable the implementation of personalized strategies for risk assessment and therapeutic intervention, ultimately improving patient outcomes. However, our study still has certain limitations. Our prognostic model needs to be validated in larger patient cohorts.

Abbreviations

COVID-19, Coronavirus Disease 2019; ACE2, angiotensin-converting enzyme 2; TLR-7, Toll-like receptor 7; SARS, Severe acute respiratory syndrome; RSF, random survival forest; Enet, elastic network; plsRcox, partial least squares regression for Cox; C-index, concordance index; CRM, COVID-19-related prognostic model; OS, overall survival; PCs, principal components; GEO, Gene expression omnibus; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; ROC, Receiver operating characteristic; AUC, Area under the curve.

Data Sharing Statement

Single-cell transcriptome data and bulk transcriptome sequencing data were obtained from the GEO database, GEO registration numbers are GSE217948, GSE158055 datasets, GSE217948, GSE179850. The code that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

All patients in this study are in compliance with the regulations and guidelines of Research Ethics Committee of the second hospital of Hebei Medical University(approval number 2023-R146).All patients in this study are in accordance with the Declaration of Helsinki ethical principles. Informed consent was obtained for publication of patient data.

Consent for publication

All authors have agreed to publish this manuscript.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically

reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure
The authors declare no conflicts of interest regarding this study.

References
1. Yüce M, Filiztekin E, Özkaya KG. COVID-19 diagnosis -A review of current methods. *Biosens Bioelectron*. 2021;172:112752. doi:10.1016/j. bios.2020.112752
2. Umakanthan S, Sahu P, Ranade AV, et al. Origin, transmission, diagnosis and management of coronavirus disease 2019 (COVID-19). *Postgrad Med J*. 2020;96(1142):753–758. doi:10.1136/postgradmedj-2020-138234
3. Ciotti M, Angeletti S, Minieri M, et al. COVID-19 Outbreak: an Overview. *Chemotherapy*. 2019;64(5–6):215–223. doi:10.1159/000507423
4. Safiabadi Tali SH, LeBlanc JJ, Sadiq Z, et al. Tools and Techniques for Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)/ COVID-19 Detection. *Clin Microbiol Rev*. 2021;34(3):e00228–20. doi:10.1128/CMR.00228-20
5. Majumder J, Minko T. Recent Developments on Therapeutic and Diagnostic Approaches for COVID-19. *AAPS J*. 2021;23(1):14. doi:10.1208/ s12248-020-00532-2
6. Soiza RL, Scicluna C, Thomson EC. Efficacy and safety of COVID-19 vaccines in older people. *Age Ageing*. 2021;50(2):279–283. doi:10.1093/ ageing/afaa274
7. Lopez Bernal J, Andrews N, Gower C, et al. Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ*. 2021;373:n1088. doi:10.1136/ bmj.n1088
8. Y-D L, Chi W-Y, J-H S, Ferrall L, Hung C-F, T-C W. Coronavirus vaccine development: from SARS and MERS to COVID-19. *J Biomed Sci*. 2020;27(1):104. doi:10.1186/s12929-020-00695-2
9. Ita K. Coronavirus Disease (COVID-19): current Status and Prospects for Drug and Vaccine Development. *Arch Med Res*. 2021;52(1):15–24. doi:10.1016/j.arcmed.2020.09.010
10. Zhang H-P, Sun Y-L, Wang Y-F, et al. Recent developments in the immunopathology of COVID −19. *Allergy*. 2023;78(2):369–388. doi:10.1111/ all.15593
11. Chilamakuri R, Agarwal S. COVID-19: characteristics and Therapeutics. *Cells*. 2021;10(2):206. doi:10.3390/cells10020206
12. Gusev E, Sarapultsev A, Solomatina L, Chereshnev V. SARS-CoV-2-Specific Immune Response and the Pathogenesis of COVID-19. *Int J Mol Sci*. 2022;23(3):1716. doi:10.3390/ijms23031716
13. Khalifa AE, Ghoneim AI. Potential value of pharmacological agents acting on toll-like receptor (TLR) 7 and/or TLR8 in COVID-19. *Curr Res Pharmacol Drug Discov*. 2021;2:100068. doi:10.1016/j.crphar.2021.100068
14. Mohamadian M, Chiti H, Shoghli A, Biglari S, Parsamanesh N, Esmaeilzadeh A. COVID-19: virology, biology and novel laboratory diagnosis. *J Gene Med*. 2021;23(2):e3303. doi:10.1002/jgm.3303
15. Brant AC, Tian W, Majerciak V, Yang W, Zheng Z-M. SARS-CoV-2: from its discovery to genome structure, transcription, and replication. *Cell Biosci*. 2021;11(1):136. doi:10.1186/s13578-021-00643-z
16. Yan L, Ge J, Zheng L, et al. Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis. *Cell*. 2021;184(1):184–193.e10. doi:10.1016/j.cell.2020.11.016
17. Ortelli P, Ferrazzoli D, Sebastianelli L, et al. Neuropsychological and neurophysiological correlates of fatigue in post-acute patients with neurological manifestations of COVID-19: insights into a challenging symptom. *J Neurol Sci*. 2021;420:117271. doi:10.1016/j.jns.2020.117271
18. Ceban F, Ling S, Lui LMW, et al. Fatigue and cognitive impairment in Post-COVID-19 Syndrome: a systematic review and meta-analysis. *Brain Behav Immun*. 2022;101:93–135. doi:10.1016/j.bbi.2021.12.020
19. Xu G, Qi F, Li H, et al. The differential immune responses to COVID-19 in peripheral and lung revealed by single-cell RNA sequencing. *Cell Discov*. 2020;6(1):73. doi:10.1038/s41421-020-00225-2
20. Shekhawat J, Gauba K, Gupta S, et al. Interleukin-6 Perpetrator of the COVID-19 Cytokine Storm. *Indian J Clin Biochem*. 2021;36(4):440–450. doi:10.1007/s12291-021-00989-8
21. Bülow Anderberg S, Luther T, Berglund M, et al. Increased levels of plasma cytokines and correlations to organ failure and 30-day mortality in critically ill Covid-19 patients. *Cytokine*. 2021;138:155389. doi:10.1016/j.cyto.2020.155389
22. Herbein G. An epigenetic signature to fight COVID-19. *EBioMedicine*. 2021;67:103385. doi:10.1016/j.ebiom.2021.103385
23. Bose M, Mitra B, Mukherjee P. Mucin signature as a potential tool to predict susceptibility to COVID-19. *Physiol Rep*. 2021;9(1):e14701. doi:10.14814/phy2.14701
24. Cappuccio A, Chawla DG, Chen X, et al. Multi-objective optimization identifies a specific and interpretable COVID-19 host response signature. *Cell Syst*. 2022;13(12):989–1001.e8. doi:10.1016/j.cels.2022.11.008
25. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: relationship to Negative RT-PCR Testing. *Radiology*. 2020;296(2):E41–5. doi:10.1148/radiol.2020200343
26. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. 2019;16 (12):1289–1296. doi:10.1038/s41592-019-0619-0
27. Zhang X, Lan Y, Xu J, et al. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res*. 2019;47(D1):D721– 8. doi:10.1093/nar/gky900

28. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 2021;12(1):1088. doi:10.1038/s41467-021-21246-9

29. Montazersaheb S, Hosseiniyan Khatibi SM, Hejazi MS, et al. COVID-19 infection: an overview on cytokine storm and related interventions. *Virol J*. 2022;19(1):92. doi:10.1186/s12985-022-01814-1

30. Yongzhi X. COVID-19-associated cytokine storm syndrome and diagnostic principles: an old and new Issue. *Emerg Microbes Infect*. 2021;10(1):266–276. doi:10.1080/22221751.2021.1884503

31. McConnell MJ, Kawaguchi N, Kondo R, et al. Liver injury in COVID-19 and IL-6 trans-signaling-induced endotheliopathy. *J Hepatol*. 2021;75(3):647–658. doi:10.1016/j.jhep.2021.04.050

32. Zizzo G, Tamburello A, Castelnovo L, et al. Immunotherapy of COVID-19: inside and Beyond IL-6 Signalling. *Front Immunol*. 2022;13:795315. doi:10.3389/fimmu.2022.795315

33. Coomes EA, Haghbayan H. Interleukin-6 in Covid-19: a systematic review and meta-analysis. *Rev Med Virol*. 2020;30(6):1–9. doi:10.1002/rmv.2141

34. Schultheiß C, Willscher E, Paschold L, et al. The IL-1β, IL-6, and TNF cytokine triad is associated with post-acute sequelae of COVID-19. *Cell Rep Med*. 2022;3(6):100663. doi:10.1016/j.xcrm.2022.100663

35. Vora SM, Lieberman J, Wu H. Inflammasome activation at the crux of severe COVID-19. *Nat Rev Immunol*. 2021;21:694–703. doi:10.1038/s41577-021-00588-x

36. Tharmarajah E, Buazon A, Patel V, et al. IL-6 inhibition in the treatment of COVID-19: a meta-analysis and meta-regression. *J Infect*. 2021;82:178–185. doi:10.1016/j.jinf.2021.03.008

37. Karki R, Sharma BR, Tuladhar S, et al. Synergism of TNF-α and IFN-γ Triggers Inflammatory Cell Death, Tissue Damage, and Mortality in SARS-CoV-2 Infection and Cytokine Shock Syndromes. *Cell*. 2021;184(1):149–168.e17. doi:10.1016/j.cell.2020.11.025

38. Mattiola I, Tomay F, De Pizzol M, et al. The macrophage tetraspan MS4A4A enhances dectin-1-dependent NK cell-mediated resistance to metastasis. *Nat Immunol*. 2019;20(8):1012–1022. doi:10.1038/s41590-019-0417-y

39. Silva-Gomes R, Mapelli SN, Boutet M-A, et al. Differential expression and regulation of MS4A family members in myeloid cells in physiological and pathological conditions. *J Leukoc Biol*. 2022;111(4):817–836. doi:10.1002/JLB.2A0421-200R

40. Zhang J, Zhang L. Bioinformatics approach to identify the influences of SARS-COV2 infections on atherosclerosis. *Front Cardiovasc Med*. 2022;9:907665. doi:10.3389/fcvm.2022.907665

41. Hu L, Zhang P, Sun W, Zhou L, Chu Q, Chen Y. PDPN is a prognostic biomarker and correlated with immune infiltrating in gastric cancer. *Medicine*. 2020;99(e19957).

42. Fassan M, Collesei A, Angerilli V, et al. Multi-Design Differential Expression Profiling of COVID-19 Lung Autopsy Specimens Reveals Significantly Deregulated Inflammatory Pathways and SFTPC Impaired Transcription. *Cells*. 2022;11(6):1011. doi:10.3390/cells11061011

43. Saheb Sharif-Askari F, Saheb Sharif-Askari N, Hafezi S, et al. Upregulation of interleukin-19 in saliva of patients with COVID-19. *Sci Rep*. 2022;12(1):16019. doi:10.1038/s41598-022-20087-w

44. Ebenig A, Muraleedharan S, Kazmierski J, et al. Vaccine-associated enhanced respiratory pathology in COVID-19 hamsters after TH2-biased immunization. *Cell Rep*. 2022;40(7):111214. doi:10.1016/j.celrep.2022.111214

45. Leigh T, Scalia RG, Autieri MV. Resolution of inflammation in immune and nonimmune cells by interleukin-19. *Am J Physiol Cell Physiol*. 2020;319(3):C457–64. doi:10.1152/ajpcell.00247.2020

46. Scott DW, Patel RP. Targeting endothelial adhesion molecule mRNA to control inflammation: novel insights into potential anti-inflammatory effects of IL-19. Focus on "Interleukin-19 decreases leukocyte-endothelial cell interactions by reduction in endothelial cell adhesion molecule mRNA stability". *Am J Physiol Cell Physiol*. 2013;305(3):C253–254. doi:10.1152/ajpcell.00120.2013

47. Ellison S, Gabunia K, Richards JM, et al. IL-19 reduces ligation-mediated neointimal hyperplasia by reducing vascular smooth muscle cell activation. *Am J Pathol*. 2014;184(7):2134–2143. doi:10.1016/j.ajpath.2014.04.001

48. Kako F, Gabunia K, Ray M, et al. Interleukin-19 induces angiogenesis in the absence of hypoxia by direct and indirect immune mechanisms. *Am J Physiol Cell Physiol*. 2016;310(11):C931–941. doi:10.1152/ajpcell.00006.2016

49. Neumann J, Prezzemolo T, Vanderbeke L, et al. Increased IL-10-producing regulatory T cells are characteristic of severe cases of COVID-19. *Clin Transl Immunology*. 2020;9(e1204). doi:10.1002/cti2.1204

50. Luo X-H, Zhu Y, Mao J, Du R-C. T cell immunobiology and cytokine storm of COVID-19. *Scand J Immunol*. 2021;93(3):e12989. doi:10.1111/sji.12989

51. Hasanvand A. COVID-19 and the role of cytokines in this disease. *Inflammopharmacology*. 2022;30(3):789–798. doi:10.1007/s10787-022-00992-2

52. Dhar SK, V K, Damodar S, Gujar S, Das M. IL-6 and IL-10 as predictors of disease severity in COVID-19 patients: results from meta-analysis and regression. *Heliyon*. 2021;7(2):e06155. doi:10.1016/j.heliyon.2021.e06155

53. Silva MJA, Ribeiro LR, Lima KVB, Lima LNGC. Adaptive immunity to SARS-CoV-2 infection: a systematic review. *Front Immunol*. 2022;13:1001198. doi:10.3389/fimmu.2022.1001198

54. Mulder K, Patel AA, Kong WT, et al. Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. *Immunity*. 2021;54(8):1883–1900.e5. doi:10.1016/j.immuni.2021.07.007

55. André S, Picard M, Cezar R, et al. T cell apoptosis characterizes severe Covid-19 disease. *Cell Death Differ*. 2022;29(8):1486–1499. doi:10.1038/s41418-022-00936-x

56. Rojas LA, Sethna Z, Soares KC, et al. Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer. *Nature*. 2023;618(7963):144–150. doi:10.1038/s41586-023-06063-y

57. Vaz de Paula CB, Nagashima S, Liberalesso V, et al. COVID-19: immunohistochemical Analysis of TGF-β Signaling Pathways in Pulmonary Fibrosis. *Int J Mol Sci*. 2021;23(1):168. doi:10.3390/ijms23010168

58. Gedda MR, Danaher P, Shao L, et al. Longitudinal transcriptional analysis of peripheral blood leukocytes in COVID-19 convalescent donors. *J Transl Med*. 2022;20:587. doi:10.1186/s12967-022-03751-7

59. Wang E-Y, Chen H, Sun B-Q, et al. Serum levels of the IgA isotype switch factor TGF-β1 are elevated in patients with COVID-19. *FEBS Lett*. 2021;595(13):1819–1824. doi:10.1002/1873-3468.14104

60. Xu J, Xu X, Jiang L, Dua K, Hansbro PM, Liu G. SARS-CoV-2 induces transcriptional signatures in human lung epithelial cells that promote lung fibrosis. *Respir Res*. 2020;21(1):182. doi:10.1186/s12931-020-01445-6