# scientific reports

Check for updates

OPEN

# Prospective multicenter study identifying prognostic biomarkers and microbial profiles in severe CAP using BALF, blood mNGS, and PBMC transcriptomics

Wanmei Song[1,2,3,5], Qingyuan Yang[1,2,3,5], Hui Lv[4,5], Yingqi Lv[1,2,3], Yanshan Jiang[1,2,3], Jieming Qu[1,2,3] & Yanan Li[1,2,3✉]

To identify potential biomarkers associated with the prognosis and severity of severe community-acquired pneumonia (SCAP), we conducted a multi-center prospective study from January 1, 2022, to December 31, 2023, enrolling 14 mild CAP and 75 SCAP patients in Shanghai, China. Patients underwent bronchoalveolar lavage fluid (BALF) and blood metagenomic next-generation sequencing (mNGS) as well as peripheral blood mononuclear cell (PBMC) transcriptomics. Among the 75 SCAP patients, 32 died within 30 days, with older age, a history of allergies, and comorbidities like cerebrovascular disease linked to worse outcomes. BALF mNGS showed greater microbial diversity, revealing a higher prevalence of pathogens, including *Acinetobacter baumannii*, *Klebsiella pneumoniae*, and *Candida albicans*, compared to mild CAP patients. RNA sequencing identified 431 differentially expressed genes in deceased SCAP patients, with significant alterations in immune pathways. Notably, microbial markers such as *Pneumocystis jirovecii* and viral markers like Human cytomegalovirus were associated with poor outcomes. Transcriptomic biomarkers, including otoferlin (OTOF), MS4A4A, and SIGLEC1, were identified as potential prognostic indicators for SCAP severity. GSEA and traditional GO/KEGG analyses identified key immune and metabolic pathways in SCAP (death) patients, including upregulation of complement activation, oxidative phosphorylation, nitrogen metabolism, while downregulation of adaptive immune response, hematopoietic cell lineage, and antigen processing pathways. These findings underscore the superiority of BALF mNGS over blood mNGS for pathogen detection, revealing microbial differences and transcriptomic alterations between mild and severe cases while identifying prognostic markers for SCAP treatment strategies.

**Keywords** SCAP, mNGS, Transcriptomics, Prognosis, Severity, Biomarkers

**Abbreviations**

| | |
|---|---|
| CAP | Community-acquired pneumonia |
| SCAP | Severe community-acquired pneumonia |
| mNGS | Metagenomic next-generation sequencing |
| ALT | Alanine aminotransferase |
| AST | Aspartate aminotransferase |
| CRP | C-reactive protein |
| FDP | Fibrin degradation products |
| HCMV | Human cytomegalovirus |
| MS4A4A | Membrane-Spanning 4-Domains, Subfamily A, Member 4A |

[1]Department of Pulmonary and Critical Care Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. [2]Institute of Respiratory Diseases, Shanghai Jiao Tong University School of Medicine, Shanghai, China. [3]Shanghai Key Laboratory of Emergency Prevention, Diagnosis and Treatment of Respiratory Infectious Diseases, Shanghai, China. [4]Department of Critical Care Medicine, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 201620, China. [5]Wanmei Song, Qingyuan yang and Hui Lv contributed equally to this work. ✉email: liyn1005@163.com

| OTOF | Otoferlin |
|---|---|
| PRO-BNP | N-terminal pro B-type natriuretic peptide |
| SIGLEC1 | Sialic acid-binding immunoglobulin-like lectin 1 |
| TTVs | Torque Teno Viruses |
| WBC | White blood cell |
| CMTs | Conventional microbiological tests |
| BALF | Bronchoalveolar lavage fluid |
| PBMC | Peripheral blood mononuclear cells |
| DEGs | Differentially expressed genes |
| NAAT | Nucleic acid amplification testing |
| ABG | Arterial blood gas |
| NIPPV | Non-invasive positive pressure ventilation |
| FOLR3 | Folate receptor 3 |
| ITGA7 | Integrin subunit alpha 7 |
| CXCL10 | C-X-C motif chemokine ligand 10 |
| CCNA1 | Cyclin A1 |
| BP | Biological process |
| CC | Cellular component |
| MF | Molecular function |
| PyV-2 | Human polyomavirus type 2 |
| IM | Infectious mononucleosis |
| SOFA | Sequential organ failure assessment |

Community-acquired pneumonia (CAP) is a globally prevalent respiratory infectious disease, and remains a major cause of morbidity and mortality worldwide[1,2]. The annual incidence of CAP is approximately 649 to 847 adults per 100,000 population, leading to around 1.6 million hospitalizations every year in United States[2]. Severe CAP (SCAP) is a common disease in intensive care unit (ICU) accounting for 17–21% of hospitalized CAP patients, and the 30-day or 6-month case fatality rates of SCAP were as high as 27% and 39%[3,4]. The timely and accurate identification of pathogens is critical for a proper diagnosis and effective treatment of CAP especially in SCAP cases. Currently, traditional etiological detection methods, such as culture, smear microscopy, antigen or antibody testing, and nucleic acid amplification testing (NAAT) are commonly used for clinical diagnosis of CAP[5]. However, the total pathogens detection rate of CAP was less than 50% when only using these routine tests[5,6]. The unknown etiology of CAP may lead to delayed or insufficient anti-infection therapy, worsening the prognosis. Therefore, early, rapid, and accurate pathogenic diagnostics tools are necessary for SCAP.

In recent years, an unbiased and detailed high-throughput nucleic acid sequencing technology, called metagenomic next-generation sequencing (mNGS) has been widely adopted in clinical settings for pathogen detection[7,8]. mNGS offers higher efficiency, increased sensitivity, faster turnaround time, and a broader pathogen spectrum including bacteria, fungi, viruses, tuberculosis, parasites and atypical pathogens[7,9]. Studies have shown that compared to conventional microbiological tests (CMTs) alone, mNGS combined with CMTs reduced the time to clinical improvement for SCAP[8]. Many studies have confirmed that mNGS is both feasible and effective in SCAP patients[7,10,11]. For example, mNGS of bronchoalveolar lavage fluid (BALF) samples from SCAP performed significantly higher sensitivity (97.54% vs 28.68%), coincidence (90.34% vs 35.17%), and negative predictive value (80.00% vs 13.21%) than CMTs[7]. However, few studies have compared the diagnostic efficiency and pathogenic spectrum of BALF mNGS, blood mNGS, CMTs together in SCAP patients[11–13]. In addition, efficiency comparison of mNGS in pathogen detection among mild CAP, SCAP (survivor), SCAP (death) has rarely been reported. Moreover, whether SCAP (death) have a specific pathogen profile contributing to worse prognosis deserves further investigation.

Nowadays, the CRUB-65 and PSI scoring system are widely used to evaluate the severity of CAP, which is helpful for predicting patients' prognosis and future surviving state[4,5]. Although it is relatively easy to distinguish mild CAP and SCAP cases by clinical scoring system, the outcome of SCAP patients is still difficult to predict due to the similarity of clinical symptoms and laboratory indicators[14]. Thus, it is crucial to further understand heterogeneity in SCAP patients and identify better prognostic biomarkers. Bulk RNA-seq of PBMC and bioinformatic analysis enable dynamic tracing of the immune system and inflammatory host responses in SCAP patients[14,15]. The construction of key gene regulatory networks based on RNA-sequencing data is highly suitable for identifying biomarkers of disease severity and prognosis[14,16]. Most previous studies have concentrated on identifying biomarkers for mild or severe CAP, but there has been limited investigation into the host PBMC transcriptome signature associated with poor outcomes in the SCAP population, with the notable exception of studies on severe COVID-19 cases during the pandemic[14,17].

This study seeks to elucidate a comprehensive landscape encompassing the microbiome, transcriptome, clinical characteristics, and laboratory markers in SCAP, with the objective of identifying novel biomarkers associated with disease prognosis and severity in SCAP patients. Therefore, we conducted a prospective multi-centers study to explore the pathogen distribution in mild CAP, SCAP (survivor), SCAP (death) patients using BALF and PBMC DNA + RNA mNGS as well as CMTs. Additionally, we further analyzed the diagnostic performance of blood or BALF DNA + RNA mNGS in CAP cases with different outcome or severity of illness. The PBMC samples of 89 patients were collected for bulk RNA-seq and host gene network changes were analyzed using bioinformatics.

## Methods

### Study design and participant enrollment

During the COVID-19 management period, a prospective multi-center study was conducted in 8 hospitals in Shanghai from January 1, 2022, to December 31, 2023, to evaluate the clinical, microbiological, and transcriptomic characteristics of mild CAP and SCAP. We obtained written informed consent from participants or their legal representative within 24 h of hospital admission. This study was approved by the Ruijin Hospital Ethics Committee, Shanghai Jiao Tong University School of Medicine (Ethics Approval No.: Ruijin Hospital Ethics Committee 2021–73); Shanghai General Hospital, Shanghai Jiao Tong University (Ethics Approval No.: 2021KY041); Zhongshan Hospital, Fudan University (Ethics Approval No.: B2021-183R); Xinhua Hospital, Shanghai Jiao Tong University School of Medicine (Ethics Approval No.: XHEC-C-2021-038-1); Shanghai Chest Hospital, Shanghai Jiao Tong University (Ethics Approval No.: 2021-130); Shanghai Public Health Clinical Center (Ethics Approval No.: 2021-S035-02); Renji Hospital, Shanghai Jiao Tong University School of Medicine (Ethics Approval No.: KY2021-103-B); and Shanghai Sixth People's Hospital, Shanghai Jiao Tong University (Ethics Approval No.: IS2147). Inclusion criteria included: adults over 18 years of age; no gender limitations; patients with a diagnosis of CAP or SCAP; and newly diagnosed pneumonia patients. Exclusion criteria included: being under 18 years of age; active infection with SARS-CoV-2 (confirmed by routine PCR or nasopharyngeal swab at admission) or diagnosed with COVID-19 pneumonia; immunosuppressed status due to other diseases or treatments, such as HIV infection or ongoing immunosuppressive therapy (e.g., organ transplant recipients, long-term use of corticosteroids or other immunosuppressive medications), which results in severe and sustained immunodeficiency; active massive hemoptysis, severe bleeding disorders, and coagulation disorders; hospitalized for more than two consecutive days in the past three months; and refusal to undergo BAL.

CAP was defined as the presence of one or more acute clinical symptoms or signs compatible with CAP, including dyspnea, cough, purulent sputum, inspiratory crackles, hypoxia, fever ($>38$ °C), and radiologic features of pneumonia on CT or X-ray. According to the guidelines of the American Thoracic Society and Infectious Diseases Society of America, SCAP was defined as CAP patients who met either one major criterion or three or more minor criteria. The criteria for SCAP diagnosis include a respiratory rate greater than 30 breaths per minute, a PaO2/FIO2 ratio of less than 250, and the presence of multilobar infiltrates. Additionally, patients may exhibit confusion or disorientation, uremia (defined as a blood urea nitrogen level exceeding 20 mg/dl), leukopenia, thrombocytopenia, hypothermia, and hypotension that necessitates aggressive fluid resuscitation. Major diagnostic criteria include septic shock with the requirement for vasopressor therapy and respiratory failure requiring mechanical ventilation. Mild community-acquired pneumonia (mild CAP) refers to patients who meet the general criteria for CAP but do not fulfill the diagnostic criteria for SCAP. These patients exhibit typical CAP symptoms, such as cough, dyspnea, and fever, but do not require intensive interventions like mechanical ventilation or vasopressor therapy, which are characteristic of SCAP.

In total, our study enrolled 89 CAP patients, including 14 mild CAP and 75 SCAP cases. All patients received treatment according to the CAP consensus guidelines in China. PBMC samples from all patients was collected within 48 h of admission and sent to the Beijing Genomics Institute for bulk RNA-seq. At the same time, BALF and PBMC samples from 14 mild CAP and 65 SCAP patients were also collected for DNA + RNA mNGS. Baseline data at admission were collected, including demographic characteristics (sex, age, BMI, smoking or drinking history), medical history, comorbidities, laboratory indicators (blood routine, blood biochemistry, electrolytes, arterial blood gas analysis, immune indicators), microbiologic results of CMTs, and all variables required for calculating CRUB-65 and PSI scores. The primary outcome parameter was 30-day all-cause mortality.

### Conventional microbiological tests (CMTs)

Bacterial and fungal cultures, microscopic smears of respiratory specimens, and polymerase chain reaction (PCR) assays of blood and throat swabs for virus detection (including influenza virus, parainfluenza virus, coronavirus, respiratory syncytial virus, adenovirus, human metapneumovirus, rhinovirus and enterovirus) were routinely conducted. Urine antigen tests (for *Streptococcus pneumoniae* and *Legionella pneumophila*) and serology tests (including β-D-1,3 glucan detection, aspergillus galactomannan detection, and pathogen specific antibody detection) were also performed after admission.

### Sample collection

BALF samples were collected from the most severe lesion sites through bedside fibreoptic bronchoscopy. Each BALF sample was divided into three tubes, two were sent to Beijing Genomics Institute for DNA and RNA mNGS, and one was used for bacterial and fungal culture in the hospital's microbiological laboratory. Disposable sterile needles were used to collect whole blood samples into three specialized nucleic acid-free tube (10 ml for each) within 48 h of admission. Specimen quality should be strictly controlled, and hemolysis was meticulously avoided. Two blood samples were sent for DNA and RNA mNGS, while the third was used for PBMC separation through density gradient centrifugation. Subsequently, 1 ml TRIzol reagent was added to each PBMC sample, which was then refrigerated at -80℃ before being sent to Beijing Genomics Institute for bulk RNA-seq.

### DNA + RNA mNGS of BALF and blood

The process of DNA and RNA mNGS predominantly involves the following steps: DNA extraction, RNA extraction, metagenomic library preparation, sequencing, and subsequent data analysis. All samples followed a standardized protocol: 300 μL for plasma and 450 μL for BALF, with 300 μL used after lysis. Quality control ensured consistent DNA/RNA extraction and minimized dilution effects. The details were as follows: (1) DNA extraction: According to the manufacturer's instructions, we extracted DNA from BALF and Blood samples using QIAamp® UCP Pathogen DNA Kit (Qiagen). Benzonase (Qiagen) and Tween20 were used to remove human DNA. (2) RNA extraction: QIAamp®Viral RNA Kit (Qiagen) and Ribo-Zero rRNA Removal Kit (Illumina) were

used to extract RNA and remove ribosomal RNA. (3) Metagenomic library preparation and sequencing: cDNA was generated from an RNA template by reverse transcription before library preparation. Then, DNA and cDNA were used separately for subsequent library construction using a Nextera XT DNA Library Prep Kit (Illumina, San Diego, CA). All libraries were quantified by Agilent 2100 (Agilent Technologies, Santa Clara, CA) and Qubit 2.0 (Invitrogen, USA). The qualified DNA or cDNA library was transformed into single-stranded circular DNA library by DNA denaturation and cyclization. Single-stranded circular DNA molecules are replicated by rolling rings to form a DNA nanosphere (DNB) containing multiple copies. The DNBs obtained were added to the mesh holes on the chip using high-density DNA nanochip technology, and sequenced by combined probe anchoring polymerization (cPAS) on MGISEQ-2000 platform (MGI, China). (4) Data analysis: First, raw data were quality-filtered using SOAPnuke (v1.5.0), and adapter sequences, low-quality reads (Q < 30), short reads (length < 36 bp), duplicate reads, and low-complexity reads were removed for subsequent analysis. Trimmed reads were aligned against the human reference genome (hs37d5) using Bowtie2 software to remove human reads. The remaining non-human sequences were then aligned to microbial genome databases of bacteria, viruses, fungi, parasites, and specific pathogens (Genome Database: http://ftp.ncbi.nlm.nih.gov/genomes/genbank/). Species annotation of the mapped data was conducted by Kraken2, the species-level abundance was estimated by Bracken using Bayesian algorithms and Kraken classification results. Finally, suspected pathogens were listed including strictly mapped read counts, coverage rates, and depths.

In each mNGS sequencing batch, two negative controls (sterile water) were included to monitor contamination, and bioinformatics pipelines were used to remove common contaminants. Strict aseptic procedures and dual-index barcoding minimized cross-contamination. Sequencing was performed on the MGI2000 platform, processing 48 samples per chip with a depth of 40–100 M reads per sample. Quality control was implemented at multiple stages: reference materials with known pathogens were used during reagent production, and bacterial-spiked positive controls were included in routine testing to ensure accuracy.

### Pathogen classification and mNGS vs. CMT comparison
Pneumonia-associated microorganisms, referred to as pathogens, are defined as microorganisms that can contribute to the onset or exacerbation of disease under specific clinical conditions. While *Corynebacterium*[18], *Enterococcus*[19], TTV[20], and *Candida*[21] are typically considered part of the normal flora in healthy individuals, they may act as opportunistic pathogens in patients with severe pneumonia, particularly those with compromised immune systems or lung tissue damage. In clinical reports provided by BGI, these microorganisms were identified as pneumonia-associated microorganisms rather than background flora due to their potential role in disease progression. The classification of background microorganisms takes into account microbial colonization characteristics in specific body sites, such as the skin, oral cavity, and gut, as well as their pathogenicity and virulence. Microorganisms that are commonly present, have stable abundance, and are not typically associated with disease in healthy individuals are considered background microorganisms. In contrast, pathogens are identified based on their potential to cause disease under certain clinical conditions. This classification, guided by mNGS technology, epidemiological data, and clinical experience, ensures that relevant microorganisms are appropriately categorized in the context of severe pneumonia. We compared CMT and mNGS results based on clinical judgment. Physicians integrated clinical, laboratory, and imaging data for pathogen identification. While results may differ, the two methods are complementary, with final identification determined by physician expertise.

### RNA-seq of PBMC and bioinformatics analysis
Initially, total RNA was extracted from PBMC of each CAP patient using TRIzol (Invitrogen). The quality, integrity, and concentration of RNA were assessed. A transcriptome library was constructed from 2 μg of RNA using KAPA RNA HyperPrep with Ribo-Zero rRNA Removal Kit (Roche). RNA-sequencing was then performed on the DNBSEQ platform to obtain raw data. Raw data were filtered using SOAPnuke (v1.5.6) to produce clean data. The following reads were removed: (1) reads containing connector sequences; (2) reads with unknown base N content (≥ 5%); (3) low-quality reads (i.e., those with a base quality score of less than 15 accounting for more than 20% of the total base number).

Clean data were compared to the reference genome using the HISAT2 (v2.1.0) and Bowtie2(v2.3.4.3). Gene expression quantification was performed using RSEM (v1.3.1), and the clustering heatmap of gene expression across different samples was generated using pheatmap (v1.0.8). Read count normalization and differential gene expression analysis (SCAP_survivor vs mild CAP, SCAP_death vs mild CAP, SCAP_survivor vs SCAP_death) were performed using DESeq2 (v1.4.5) with a Q value ≤ 0.05 or FDR ≤ 0.001. Specifically, comparisons between Mild CAP and SCAP (Survivor), as well as between Mild CAP and SCAP (Death), were conducted to assess the transcriptomic differences between mild and severe pneumonia. Additionally, a comparison between SCAP (Survivor) and SCAP (Death) was performed to investigate gene expression patterns associated with patient prognosis. GO and KEGG pathway enrichment analyses of differentially expressed genes (DEGs) were performed using Phyper, with a threshold is set to Qvalue ≤ 0.05. The pathway analysis was performed using the KEGG database[22–24].

### Statistical analysis
Continuous variables conforming to a normal distribution were presented as the mean ± standard deviation (x ± s), and an unpaired t-test was used for group comparisons. Clinical indicators that exhibited a skewed distribution were reported as the median with the interquartile range [M (P25, P75)], and the groups were compared using the Mann–Whitney U test. Categorical variables were represented as absolute numbers and percentages [N(%)] and were compared using the Chi-square test or Fisher's exact test. Statistical analysis and plotting were performed using SPSS 22.0 and Graphpad Prism 8 software. A *P* value of ≤ 0.05 was considered statistically significant.

## Results

### Baseline characteristics of enrolled patients

A total of 89 patients were enrolled in our study, comprising 14 with mild CAP and 75 with SCAP, of whom 32 SCAP (42.67%, 32/75) patients died during the 30-day follow-up period. Their demographic and clinical characteristics are detailed in Table 1. Patients in the SCAP (death) group were older than those in the SCAP (survivor) group (69.00 ± 15.64 vs 58.56 ± 17.57, $P = 0.0095$). Compared to mild CAP patients, a higher proportion of males was observed in both the SCAP (survivor) (86.05% vs 42.86%, $P = 0.0011$) and SCAP (death) (75% vs 42.86%, $P = 0.0352$) group, although gender was not significantly correlated with SCAP prognosis. Interestingly, allergic history was more prevalent in SCAP (death) group compared to both mild CAP (28.13% vs 15.38%) and SCAP (survivor) patients (28.13% vs 9.52%, $P = 0.0372$). The incidences of comorbidities (93.75% vs 61.54% $P = 0.0069$) and cerebrovascular disease (31.25% vs 7.69%, $P = 0.0956$) were higher in SCAP (death) group than in mild CAP cases. Additionally, SCAP (death) patients had a higher rate of cerebrovascular disease (31.25% vs 7.14%, $P = 0.0022$) compared to SCAP (survivor) patients, suggesting it may be an indicator of worse outcomes in SCAP patients. Both the SCAP (survivor) (36.59% vs 0%, $P = 0.0080$) and SCAP (death) (33.33% vs 0%, $P = 0.0140$) groups were more likely to use non-invasive mechanical ventilation (NIPPV) compared to the mild CAP group.

### Laboratory findings at admission

As shown in Table 2, blood routine, blood biochemistry, blood electrolytes, arterial blood gas (ABG), coagulation indicators, serum immunoglobulin, inflammatory indicators, and other clinical laboratory indicators of CAP and SCAP patients were collected upon admission. As the severity of CAP increased [mild CAP→SCAP(survivor)→SCAP(death)], the levels of white blood cell count and neutrophil count, lactic acid in ABG, alanine aminotransferase (ALT), and aspartic transaminase (AST), CRP first increased and then decreased, while lymphocyte count and total protein (TP) continuously declined, $P < 0.05$. Additionally, laboratory indicators, including fibrin degradation product (FDP), D-dimer, PRO-BNP, CK-MB, Myoglobin, serum IgE, and β-D-1, 3-glucan (fungus) were positively correlated with the severity of CAP and poor prognosis, $P < 0.05$. SCAP patients had a lower level of eosinophil count, Ca2 + and serum IgM than mild CAP patients,

| Characteristics | Mild CAP(n = 14) | SCAP (survivor) (n = 43) | SCAP (death) (n = 32) | SCAP (survivor) vs Mild CAP | SCAP (death) vs Mild CAP | SCAP (death) vs SCAP (survivor) |
|---|---|---|---|---|---|---|
| Age (n = 14/43/32) | 62.29 ± 14.90 | 58.56 ± 17.57 | 69.00 ± 15.64 | 0.4785 | 0.1812 | 0.0095** |
| Sex (n = 14/43/32) | | | | | | |
| Male | 6(42.86%) | 37(86.05%) | 24(75%) | 0.0011** | 0.0352* | 0.2246 |
| Female | 8(57.14%) | 6(13.95%) | 8(25%) | | | |
| BMI(n = 13/38/27) | 24.39 ± 4.923 | 23.29 ± 2.957 | 22.82 ± 5.041 | 0.3388 | 0.3585 | 0.6374 |
| Previous history | | | | | | |
| Smoking (n = 13/42/32) | 6(46.15%) | 16(37.21%) | 10(31.25%) | 0.6043 | 0.3438 | 0.5411 |
| Drinking (n = 13/42/32) | 2(15.38%) | 8(18.6%) | 3(9.38%) | 0.7648 | 0.5610 | 0.2466 |
| Operation history (n = 13/42/29) | 4(30.77%) | 13(30.95%) | 9(31.03%) | 0.9900 | 0.9863 | 0.9941 |
| Allergic history (n = 13/42/32) | 2(15.38%) | 4(9.52%) | 9(28.13%) | 0.5536 | 0.3674 | 0.0372* |
| Comorbidities (n = 13/42/32) | 8(61.54%) | 36(85.71%) | 30(93.75%) | 0.0569 | 0.0069** | 0.2701 |
| Cancer (n = 13/42/32) | 1(7.69%) | 7(16.67%) | 4(12.5%) | 0.4225 | 0.6418 | 0.6177 |
| Chronic lung disease(n = 13/42/32) | 5(38.46%) | 11(26.19%) | 8(25%) | 0.3946 | 0.3665 | 0.9075 |
| Diabetes(n = 13/42/31) | 0(0%) | 7(16.67%) | 7(22.58%) | 0.1151 | 0.0617 | 0.5258 |
| Hypertension(n = 13/42/31) | 3(23.08%) | 13(30.95%) | 15(48.39%) | 0.5848 | 0.1192 | 0.1300 |
| Cardiac insufficiency(n = 13/42/31) | 2(15.38%) | 5(11.9%) | 3(9.68%) | 0.7422 | 0.5863 | 0.3011 |
| Renal insufficiency (n = 13/42/31) | 2(15.38%) | 5(11.9%) | 1(3.23%) | 0.7422 | 0.1443 | 0.1820 |
| Cerebrovascular disease(n = 13/42/32) | 1(7.69%) | 3(7.14%) | 10(31.25%) | 0.6843 | 0.0956 | 0.0022** |
| Neurological diseases (dementia, Parkinson's, etc.) (n = 13/42/31) | 0(0%) | 2(4.76%) | 4(12.9%) | 0.4228 | 0.1743 | 0.2106 |
| Mental disease (n = 13/42/30) | 0(0%) | 3(7.14%) | 2(6.67%) | 0.3217 | 0.3486 | 0.9080 |
| Respiratory therapy (n = 14/41/30) | | | | | | |
| Nasal tube oxygen | 8(57.14%) | 15(36.59%) | 15(50%) | 0.1782 | 0.6586 | 0.2583 |
| Mask oxygen inhalation | 0(0%) | 2(4.88%) | 1(3.33%) | 0.3999 | 0.4895 | 0.7493 |
| Nasal high flow humidified oxygen therapy (HFNC) | 0(0%) | 9(21.95%) | 3(10%) | 0.0553 | 0.2203 | 0.1844 |
| Non-invasive mechanical ventilation (NIPPV) | 0(0%) | 1(2.44%) | 1(3.33%) | 0.5554 | 0.4895 | 0.8220 |
| Invasive mechanical ventilation (IPP) | 0(0%) | 15(36.59%) | 10(33.33%) | 0.0080** | 0.0140* | 0.7769 |

**Table 1**. Demographic and clinical characteristics of the 89 patients enrolled with mild CAP or SCAP. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

| Laboratory indicators [median (25%, 75%)] | References | Mild CAP(n = 14) | SCAP (survivor) (n = 43) | SCAP (death) (n = 32) | SCAP (survivor) vs Mild CAP | SCAP (death) vs Mild CAP | SCAP (death) vs SCAP (survivor) |
|---|---|---|---|---|---|---|---|
| Blood routine | | | | | | | |
| Red blood cell count | $(2.3–5.8) \times 10^{12}$/L | 3.79 (3.290, 4.160) | 3.765 (3.313, 4.170) | 3.67 (3.060, 4.260) | 0.8988 | 0.652 | 0.4722 |
| Hemoglobin | 130–175 g/L | 111 (75.50, 130.5) | 111 (96.00, 128.0) | 109 (89.00, 124.0) | 0.7061 | 0.9342 | 0.5846 |
| White blood cell count | $(3.5–9.5) \times 10^9$/L | 6.93 (5.015, 10.14) | 11.81 (9.143, 16.08) | 9.28 (6.380, 13.50) | 0.0002** | 0.0948 | 0.0287* |
| Neutrophil count | $(1.8–6.3) \times 10^9$/L | 6.11 (3.890, 10.12) | 10.7 (8.200, 16.40) | 7.8 (5.240, 11.60) | 0.0068* | 0.3653 | 0.0093** |
| Lymphocyte count | $(1.1–3.2) \times 10^9$/L | 1.34 (0.820, 1.755) | 0.88 (0.475, 1.453) | 0.52 (0.290, 0.800) | 0.0532 | < 0.0001*** | 0.0065** |
| Eosinophil count | $(0.02–0.52) \times 10^9$/L | 0.145 (0.0625, 0.2175) | 0.01 (0.000, 0.110) | 0.01 (0.000, 0.050) | 0.0122* | 0.0026* | 0.4143 |
| Basophil Count | $(0.00–0.06) \times 10^9$/L | 0.03 (0.015,0.045) | 0.02 (0.000, 0.045) | 0.01 (0.000, 0.030) | 0.2482 | 0.0967 | 0.6671 |
| Platelet count | $(125–350) \times 10^9$/L | 223 (169.0, 309.0) | 204.5 (131.5, 280.5) | 173 (119.0, 249.0) | 0.4129 | 0.2033 | 0.5847 |
| Blood biochemistry | | | | | | | |
| Blood urea nitrogen | 2.5–7.1 mmol/L | 5 (3.455, 10.05) | 6.065 (4.623, 9.903) | 8.7 (5.215, 12.86) | 0.2437 | 0.1326 | 0.1763 |
| Serum creatinine | 62–115 μmol/L | 59 (53.00,87.95) | 70.2 (53.00, 107.1) | 75 (59.10, 100.4) | 0.5274 | 0.3056 | 0.6044 |
| Alkaline phosphatase (ALP) | 45–125 IU/L | 93.5 (54.25, 123.0) | 100 (72.00, 163.0) | 78 (62.00, 117.5) | 0.2881 | 0.8101 | 0.0599 |
| Alanine aminotransferase (ALT) | 9–50 IU/L | 19 (11.00, 49.50) | 45.5 (14.75, 107.8) | 26 (19.00, 56.00) | 0.0472* | 0.3168 | 0.2263 |
| Aspartic transaminase (AST) | 15–40 IU/L | 39.5 (14.50, 39.50) | 109.8 (27.00, 109.8) | 57 (22.07, 57.00) | 0.0125* | 0.1767 | 0.0601 |
| Total bilirubin (TBIL) | 4.7–24 μmol/L | 12.8 (8.700, 15.05) | 13.7 (8.800, 22.10) | 12 (9.825, 18.73) | 0.5410 | 0.7991 | 0.6672 |
| Total protein (TP) | 60–83 g/L | 67(59.94,73.00) | 58 (54.00,66.00) | 57.07 (53.00,63.00) | 0.0270* | 0.007** | 0.3570 |
| Prealbumin | 200–430 mg/L | 125 (59.50, 210.1) | 87.8 (43.75,198.8) | 89 (65.00,137.0) | 0.3312 | 0.2397 | 0.8529 |
| Albumin | 35–55 g/L | 34.1(28.60,38.35) | 33 (28.00,34.00) | 31 (27.08, 33.48) | 0.2170 | 0.1527 | 0.387 |
| Triglyceride (TG) | 0.38–1.88 mmol/L | 1.08 (0.970,1.500) | 1.67 (1.000,2.350) | 1.28 (1.075, 2.088) | 0.3162 | 0.6254 | 0.2204 |
| Total cholesterol (TC) | 3.35–5.69 mmol/L | 4.26(2.070,5.030) | 3.145 (2.820, 3.883) | 2.88 (2.153, 3.983) | 0.4910 | 0.1735 | 0.2365 |
| High density lipoprotein cholesterol (HDL-C) | 0.91–2.1 mmol/L | 0.78 (0.470,1.000) | 0.66 (0.610, 0.840) | 0.755 (0.450, 1.093) | 0.7667 | 0.9193 | 0.7242 |
| Low density lipoprotein cholesterin (LDL-C) | 0–3.33 mmol/L | 2.85(1.210,3.360) | 1.88(1.510,2.400) | 1.455 (1.013, 2.490) | 0.3339 | 0.1101 | 0.1810 |
| Blood glucose | 3.89–5.83 mmol/L | 4.7(4.190,7.400) | 7.29 (5.550,8.383) | 7.8 (5.790,10.10) | 0.1648 | 0.0867 | 0.4250 |
| Glycosylated hemoglobin (HbA1c) | | 5.65 (5.350,6.175) | 6.05 (5.700,11.88) | 6.25 (5.400,7.875) | 0.1719 | 0.381 | 0.5402 |
| Blood electrolytes | | | | | | | |
| K + | 3.50–5.10 mmol/L | 3.995(3.465,3.995) | 4.08(3.410,4.080) | 4.438(3.688,4.438) | 0.6981 | 0.0626 | 0.0768 |
| Na + | 130–147 mmol/L | 140(138.0,141.6) | 137.6(136.0, 141.8) | 139.0(135.9,144.9) | 0.4825 | 0.9363 | 0.5320 |
| Cl– | 99–110 mmol/L | 105(101.0,106.0) | 101.7(99.00, 105.3) | 103(99.90, 106.5) | 0.1344 | 0.3558 | 0.4096 |
| Ca2 + | 2.00–2.52 mmol/L | 2.22(1.950,2.310) | 1.965(1.880,2.163) | 2.02(1.960,2.110) | 0.0175* | 0.0743 | 0.242 |
| P2– | 0.85–1.51 mmol/L | 1.06(0.770,1.330) | 1.1 (0.8700,1.380) | 0.99 (0.7450,1.280) | 0.6764 | 0.5742 | 0.1876 |
| Mg2 + | 0.7—1.0 mmol/L | 0.855(0.7925,0.9600) | 0.83(0.7900,0.9500) | 0.82(0.7650,0.9650) | 0.7769 | 0.7286 | 0.8964 |
| Arterial blood gas (ABG) | | | | | | | |
| PaO2 | 83–108 mmHg | 90(81.01, 108.8) | 83.4(63.75,102.4) | 89.2(69.00,119.5) | 0.3301 | 0.8375 | 0.3474 |
| SpO | 93.0–98.0% | 96.9(95.00,99.00) | 97.35 (93.13, 98.70) | 97.6(95.70,99.23) | 0.8323 | 0.6351 | 0.3991 |
| PaCO2 | 35–45 mmHg | 39.45(35.00,41.00) | 38(32.25,40.67) | 38.22(34.25,46.25) | 0.8149 | 0.5509 | 0.4129 |
| PH | 7.35–7.45 | 7.44(7.430,7.450) | 7.43(7.373,7.475) | 7.39(7.360,7.460) | 0.6735 | 0.2014 | 0.3591 |
| BEb | − 3.0–3.0 mmol/L | 1.2(− 0.3000,2.800) | 1.33 (− 1.800,4.800) | − 0.19(− 3.100,3.900) | 0.8621 | 0.9610 | 0.5622 |
| Standard bicarbonate (SB) | 22–26 mmol/L | 26(24.75,27.03) | 24.4(22.00,28.40) | 24.3(21.70,28.00) | 0.6527 | 0.5238 | 0.7020 |
| Actual bicarbonate (AB) | 22–26 mmol/L | 25.95(24.75,27.25) | 25.9(22.40,28.40) | 27.1(22.90,28.60) | 0.8853 | 0.8644 | 0.8234 |
| Lactic acid | 0.5–1.6 mmol/L | 0.9(0.600,1.500) | 1.65(1.031,2.770) | 1.062(1.000,1.661) | 0.0109* | 0.1147 | 0.0164* |
| Coagulation indicators | | | | | | | |
| PT | 10.0–16.0 s | 12.7(12.00,17.20) | 13.6(12.70,15.20) | 13.7(12.63,16.38) | 0.6157 | 0.5126 | 0.5848 |
| INR | / | 1.08(1.040,1.740) | 1.16(1.100,1.300) | 1.155(1.073,1.395) | 0.4041 | 0.6947 | 0.8717 |
| APTT | 22.3–38.7 s | 30.2(26.50,38.40) | 30.8(27.53,36.55) | 31.9(27.03,38.88) | 0.7696 | 0.5954 | 0.7229 |
| Fibrin degradation product (FDP) | 0–5 mg/L | 3.05(2.500,6.800) | 9.7(4.615,17.05) | 14.14(7.600,25.25) | 0.0069** | 0.0009** | 0.1621 |
| D-dimer quantification | < 0.55 mg/L | 0.93(0.3250,2.503) | 2.42(1.450,4.930) | 3.42(1.890,6.450) | 0.0110* | 0.0034** | 0.3176 |
| TT | 14.00–21.00 s | 16(14.40,18.50) | 15.8(14.90,16.95) | 16.3(15.35,17.83) | 0.4406 | 0.7293 | 0.1281 |
| Fg | 1.8–3.5 g/L | 4.495(2.452,6.000) | 5.16(2.400,7.240) | 4.85(3.768,6.455) | 0.6257 | 0.2757 | 0.2757 |
| Myo-cardial enzymonram | | | | | | | |
| PRO-BNP | 5–349 pg/mL | 102(40.30,374.4) | 288(104.0,1115) | 628.3(168.8,988.7) | 0.0612 | 0.0117* | 0.3176 |
| Continued | | | | | | | |

| Laboratory indicators [median (25%, 75%)] | References | Mild CAP(n = 14) | SCAP (survivor) (n = 43) | SCAP (death) (n = 32) | SCAP (survivor) vs Mild CAP | SCAP (death) vs Mild CAP | SCAP (death) vs SCAP (survivor) |
|---|---|---|---|---|---|---|---|
| Hypersensitive troponin I(cTnT) | < 0.1 μg/L | 2.7(0.011,8.150) | 0.077(0.0145,4.308) | 5.8(0.070,32.60) | 0.7607 | 0.1118 | 0.0143* |
| Creatine kinase | 22–269 IU/L | 48.5(25.25,67.50) | 61(22.00,200.0) | 54.5(34.00,115.3) | 0.4340 | 0.3610 | 0.9597 |
| CK-MB | 0.3–4 ng/mL | 1.1(0.800,2.050) | 3.3(0.900,9.000) | 2.5(1.150,5.250) | 0.0206* | 0.0372* | 0.4187 |
| Myoglobin | 0-58 ng/mL | 21.1(15.85, 54.05) | 61.7(28.70, 236.4) | 90(28.25, 281.6) | 0.0233* | 0.0076** | 0.6020 |
| Lactic dehydrogenase (LDH) | 120–250 IU/L | 187.6(166.5,293.5) | 281(202.5,406.0) | 251.5 (199.0,396.8) | 0.0744 | 0.0515 | 0.8579 |
| Cellular immunity | | | | | | | |
| CD3^ ( +) | 47.98–82.6% | 74(61.16,81.91) | 68.03(55.34,77.95) | 66.22(56.15,73.00) | 0.3917 | 0.2327 | 0.5177 |
| CD3^ ( +) CD4^ ( +) | 24.15–51.66% | 47.65(20.06,51.53) | 38.41(34.54,48.85) | 39.2(30.40,48.35) | 0.8648 | 0.8228 | 0.6584 |
| CD3^ ( +) CD8^ ( +) | 14.24–41.48% | 30.1(16.13,41.43) | 23.94(11.43,37.95) | 23.3(16.99,26.30) | 0.5162 | 0.3124 | 0.8849 |
| CD4^ ( +) /CD8^ ( +) | 1.00–2.50 | 1.84(1.130,3.435) | 1.45(1.000,2.810) | 1.51(1.225,1.968) | 0.7427 | 0.4950 | 0.9132 |
| CD3 absolute count | 713–2368 cells/μl | 797(294.0,1488) | 555.5(416.3,919.8) | 446(270.0,608.5) | 0.6599 | 0.3193 | 0.4228 |
| CD4 absolute count | 384–1346 cells/μl | 492(161.8,1002) | 260(235.0,590.0) | 307(133.5,380.0) | 0.6116 | 0.2827 | 0.5261 |
| CD8 absolute count | 220–1110 cells/μl | 306(116.4, 451.5) | 152(120.0,237.5) | 144(61.00,229.0) | 0.3980 | 0.1893 | 0.3875 |
| Inflammatory indicators | | | | | | | |
| CRP | < 10 mg/L | 29(7.075,81.03) | 107.5(20.10,166.0) | 88.7(54.43,149.0) | 0.0683 | 0.0464* | 0.6781 |
| PT | < 0.50 ng/mL | 0.12(0.030,0.520) | 0.38(0.200,7.120) | 0.33(0.150,1.440) | 0.0276* | 0.0514 | 0.3876 |
| IL-6 | < 5.4 pg/mL | 21.1(9.225,43.79) | 36.87(4.225,113.5) | 52.55(25.95,162.2) | 0.659 | 0.1059 | 0.3696 |
| IL-1β | < 3.1 pg/mL | 7.36(2.850,19.17) | 5(2.400,8.900) | 4(2.400,5.225) | 0.5243 | 0.4959 | 0.9349 |
| IL-2 receptor | < 5.71 pg/mL | 1751(4.260,4314) | 1095(2.440,1720) | 1300(189.8,3368) | 0.2606 | 0.7394 | 0.6280 |
| IL-8 | < 20.6 pg/mL | 40.5(2.400,659.4) | 36.1(2.400,94.25) | 18.15(2.400,52.63) | 1.000 | 0.8038 | 0.7487 |
| IL-10 | < 12.9 pg/mL | 2.7(2.440, 90.00) | 5(2.440, 8.000) | 5(3.350, 13.85) | 0.8040 | 0.3946 | 0.4469 |
| TNF | < 16.5 pg/mL | 7.76(2.980,57.15) | 7.435(2.400,13.00) | 9.9(3.825,12.73) | 0.8020 | 0.9714 | 0.8495 |
| Serum immunoglobulin | | | | | | | |
| IgG | 8.6–17.4 g/L | 9.81(6.998,16.08) | 11.75(9.178,14.60) | 9.525(7.550,11.85) | 0.5495 | 0.6590 | 0.0793 |
| IgM | 1–4.2 g/L | 1.675(1.268,2.188) | 03.535(0.3625,1.528) | 0.845(0.4050,1.755) | 0.0426* | 0.1093 | 0.4928 |
| IgA | 0.3–2.2 g/L | 2.615(1.608,3.185) | 2.02(1.413,2.788) | 2.325(1.690,3.058) | 0.5728 | 0.8411 | 0.5699 |
| IgE | 5.0–165.3 IU/L | 28(10.55,56.68) | 114.8(25.30,228.8) | 178(67.40,344.0) | 0.1179 | 0.0150* | 0.3669 |
| Other indicators | | | | | | | |
| β-D-1, 3-glucan (fungus) | < 60 pg/mL | 10(10,10) | 19(10,35.90) | 28.76(10,85.79) | 0.0142* | 0.0056** | 0.1905 |
| | | 0(0%) | 4(9.3%) | 6(18.50%) | 0.2366 | 0.0823 | 0.2339 |
| GM test (Aspergillus galactomannan detection) | < 0.25 ng/mL | 0.11(0.1,0.13) | 0.1(0.0365,0.214) | 0.11(0.100,0.150) | 0.5113 | 0.9456 | 0.4262 |
| | | 0(0%) | 2(4.65%) | 3(9.25%) | 0.4114 | 0.2360 | 0.4173 |

**Table 2**. Laboratory Findings at Admission. * *P* < 0.05, ** *P* < 0.01, *** *P* < 0.001.

but this indicator had no significant association with SCAP prognosis. SCAP (death) had more hypersensitive troponin I than SCAP (survivor) group.

### Association between CAP severity and pathogen detection rates by mNGS and CMTs

BALF mNGS, blood mNGS, and CMTs were all performed in mild CAP (n = 13), SCAP (survivor) (n = 40), SCAP (death) (n = 24) to compare the detection performance (Fig. 1). The comparison of the positive rate of pathogens in mild CAP and SCAP patients using blood and BALF mNGS or CMTs is showed in Fig. 2 and Appendix Table 1. Based on BALF mNGS data, SCAP (survivor) patients had a higher positive rate of bacteria compared to both mild CAP (82.5% vs 53.85%, *P* = 0.037) and SCAP (death) patients (82.5% vs 58.33%, *P* = 0.0341). In contrast to the above results, the positive rates of bacteria detected by CMTs increased with CAP severity [mild CAP vs SCAP (survivor) vs SCAP (death): 23.08% vs 42.5% vs 62.5%]. The detection rate of fungi by CMTs (7.69% vs 20% vs 29.17%) or BALF mNGS (15.38% vs 42.5% vs 50%) also showed a gradual increase as CAP severity worsened. Furthermore, the detection rate of fungi by blood mNGS was low and showed no significant differences among the three CAP groups (7.69% vs 7.5% vs 8.33%). The positive rate of DNA viruses detected by blood mNGS increased with disease severity (61.54% vs 80% vs 91.67%), whereas the rate of DNA viruses detected by BALF mNGS only increased in SCAP (death) patients (69.23% vs 67.5% vs 87.5%). In summary, SCAP patients, particularly SCAP (death) patients had higher rates of bacteria, fungi and DNA viruses according to a comprehensive analysis of BALF mNGS, blood mNGS, and CMTs data.

### Comparison of detection performance: BALF mNGS, blood mNGS, and CMTs

As shown in Fig. 2 and Table 3, BALF mNGS significantly improved bacteria, fungi, and DNA virus detection in SCAP (survivor) patients compared to CMTs. When comparing the performance of mNGS across different sample types, we found that BALF mNGS detected more bacteria in both SCAP (survivor) (82.5% vs 47.5%,
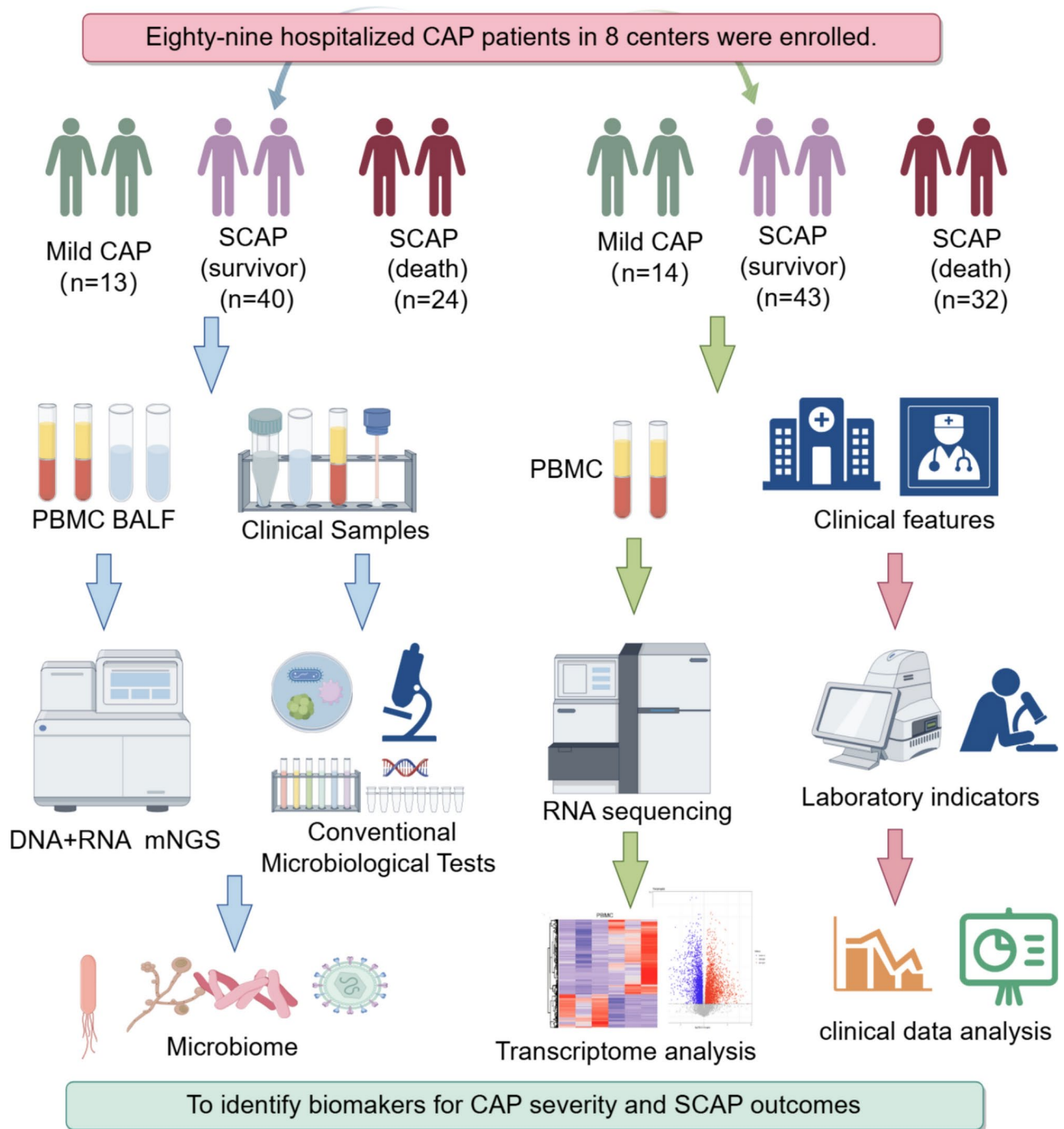
**Fig. 1**. Flow chart of the study.

$P = 0.001$) and SCAP (death) patients (58.33% vs 33.33%, $P = 0.0822$) than blood mNGS. Similarly, the detection rate of fungi by BALF mNGS was also much higher than by blood mNGS in both SCAP (survivor) (42.5% vs 7.5%, $P = 0.0003$) and SCAP (death) (50% vs 8.33%, $P = 0.0015$) groups. The detection performance of BALF mNGS on DNA virus was much better than that of CMTs, but slightly worse than blood mNGS. Blood mNGS might be more suitable for the detection of DNA viruses than BALF mNGS, particularly in SCAP (death) patients (87.5% vs 67.5%, $P = 0.0322$), but it had poor detection efficiency for fungi in all CAP patients. It is disappointing that the combined application of BALF and blood mNGS did not exhibit better performance than BALF mNGS alone for detecting bacteria and fungi among CAP or SCAP patients. In conclusion, pathogen detection rates were influenced by both the testing method and disease severity stratification. BALF mNGS and CMTs significantly improved pathogen detection rates in CAP and SCAP patients. Given the different sample sources and patient conditions, the varying detection rates between methods were expected.
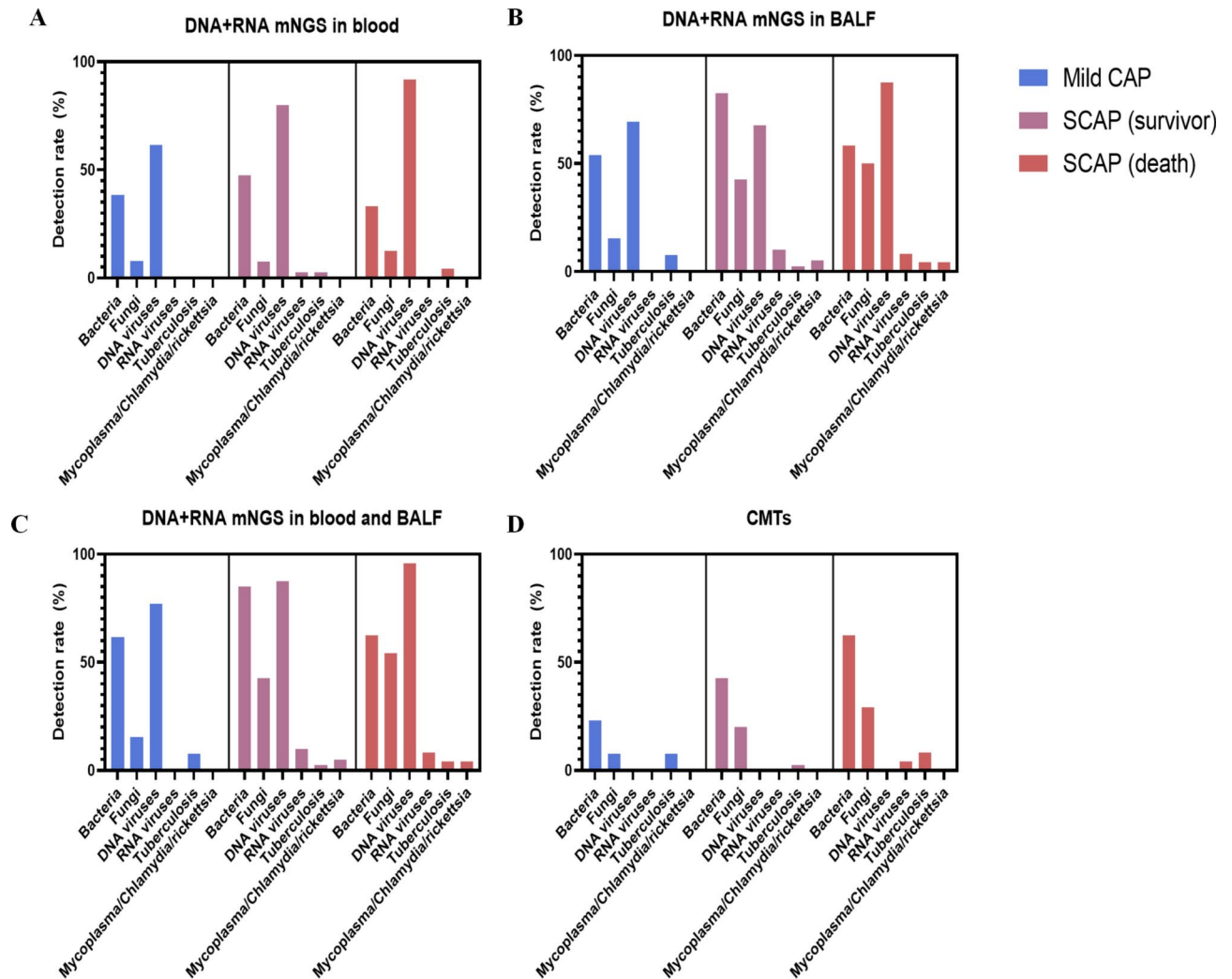
**Fig. 2.** Comparison of Pathogen Detection Rates using DNA + RNA mNGS in Blood and BALF Samples Versus CMTs in Mild and Severe CAP Cases.

### Different pathogen spectra in mild CAP, SCAP (survivor), and SCAP (death) patients

As shown in Fig. 3 and Appendix Table 2, the pathogen spectra identified in mild CAP, SCAP (survivor), and SCAP (death) patients varied across different diagnostic methods (BALF mNGS, blood mNGS, and CMTs). In mild CAP patients, BALF mNGS detected a broader range of bacteria, including *Acinetobacter baumannii* (23.08%) and *Klebsiella pneumoniae* (15.38%), compared to blood mNGS and CMTs. Among SCAP (survivor) patients, *K. pneumoniae* and *A. baumannii* were most frequently identified, with BALF mNGS showing the highest detection rates. In SCAP (death) patients, BALF mNGS revealed a higher prevalence of bacteria such as *A. baumannii* (20.83%) and *Corynebacterium striatum* (20.83%), while blood mNGS and CMTs detected *K. pneumoniae* (20.83%) and *A. baumannii* (25%) with a slightly different distribution. Fungal and viral spectra also differed by disease severity. In SCAP (survivor) patients, *Candida glabrata* (17.5%) and *Candida albicans* (15%) were the most common fungi detected by BALF mNGS, whereas in SCAP (death) patients, *C. albicans* (29.17%) and *Pneumocystis jiroveci* (16.67%) were predominant. Viral detection showed that Epstein-Barr virus (EBV) and Torque teno virus (TTVs) were the most frequent in both SCAP groups, with higher detection rates in blood mNGS compared to BALF mNGS.

DNA + RNA mNGS analysis identified significant pathogen-associated risk factors for severe CAP and poor prognosis (Table 4). In BALF samples, *P. jiroveci* (16.67% vs 2.5%, $P = 0.0409$) was more prevalent in fatal cases. Human cytomegalovirus (HCMV) in blood samples showed a higher detection rate in severe CAP cases (29.17% vs 7.69%, $P = 0.0982$). *C. striatum*, *C. albicans*, EBV, and TTVs had higher detection rates in severe cases, but these differences were not statistically significant. The results for DNA mNGS and RNA mNGS are provided separately in Appendix Table 2. Additionally, the information about co-infection conditions, read counts, and other details can be found in Appendix Tables 5–7.

| Patient type | Blood + BALF mNGS vs CMTs | Blood mNGS vs CMTs | BALF mNGS vs CMTs | Blood + BALF mNGS vs BALF mNGS | Blood + BALF mNGS vs Blood mNGS | BALF mNGS vs Blood mNGS |
|---|---|---|---|---|---|---|
| Detection rate of Bacteria | | | | | | |
| Mild CAP(n = 13) | 61.54%vs 23.08% (P = 0.0472) | 38.46% vs 23.08% (P = 0.3954) | 53.85% vs 23.08% (P = 0.2262) | 61.54% vs 53.85% (P = 0.6914) | 61.54% vs 38.46% (P = 0.2393) | 53.85% vs 38.46% (P = 0.4314) |
| SCAP (survivor) (n = 40) | 85% vs 42.5% ( P < 0.0001) | 47.5% vs 42.5% (P = 0.6531) | 82.5% vs 42.5% ( P < 0.0001) | 85% vs 82.5% (P = 0.3031) | 85% vs 47.5% (P = 0.0004) | 82.5% vs 47.5% (P = 0.001) |
| SCAP (death) (n = 24) | 62.5% vs 62.5% (P = 1.000) | 33.33% vs 62.5% (P = 0.0431) | 58.33% vs 62.5% (P = 0.7679) | 62.5% vs 58.33% (P = 0.7679) | 62.5% vs 33.33% (P = 0.0431) | 58.33% vs 33.33% (P = 0.0822) |
| Detection rate of Fungi | | | | | | |
| Mild CAP(n = 13) | 15.38% vs 7.69% (P = 0.5393) | 7.69% vs 7.69% (P = 1.000) | 15.38% vs 7.69% (P = 0.5393) | 15.38% vs 15.38% (P = 1.000) | 15.38% vs 7.69% (P = 0.5393) | 15.38% vs 7.69% (P = 0.5393) |
| SCAP (survivor) (n = 40) | 42.5% vs 20% (P = 0.0299) | 7.5% vs 20% (P = 0.1045) | 42.5% vs 20% (P = 0.0299) | 42.5% vs 42.5% (P = 1.000) | 42.5% vs 7.5% (P = 0.0003) | 42.5% vs 7.5% (P = 0.0003) |
| SCAP (death) (n = 24) | 50% vs 29.17% (P = 0.1400) | 8.33% vs 29.17% (P = 0.0645) | 50% vs 29.17% (P = 0.1400) | 50% vs 50% (P = 1.000) | 50% vs 8.33% (P = 0.0015) | 50% vs 8.33% (P = 0.0015) |
| Detection rate of DNA virus | | | | | | |
| Mild CAP(n = 13) | – | – | – | 76.92% vs 69.23% (P = 0.6584) | 76.92% vs 61.54% (P = 0.5683) | 61.54% vs 76.92% (P = 0.5683) |
| SCAP (survivor) (n = 40) | – | – | – | 87.5% vs 67.5% (P = 0.0322) | 87.5% vs 80% (P = 0.2039) | 67.5% vs 87.5% (P = 0.0322) |
| SCAP (death) (n = 24) | – | – | – | 95.83% 87.5% (P = 0.2963) | 95.83% vs 91.67% (P = 0.6366) | 87.5% vs 95.83% (P = 0.2963) |

**Table 3**. Comparison of pathogen detection performance between DNA + RNA mNGS in BALF and blood samples versus CMTs among mild CAP, SCAP (survivor), SCAP (death) patients.

### Concordance of mNGS with CMTs in mild CAP and SCAP patients

The positive and negative concordance rates between mNGS and CMTs are presented in Appendix Table 3. Concordance here refers to the agreement between mNGS and CMT in detecting pathogens. BALF mNGS exhibited a higher positive concordance rate with CMTs compared to blood mNGS across all three CAP groups, however, it demonstrated a lower negative concordance rate in SCAP patients relative to blood mNGS. The positive concordance rates of BALF mNGS with CMTs were 100%, 76.47%, 66.67% for bacterial detection, and 100%, 62.50%, 57.14% for fungal detection in mild CAP, SCAP (survivor), SCAP (death) patients, respectively. Additionally, the negative concordance rates of BALF mNGS to CMTs were 60%, 13.04%, 55.56% for bacterial detection and 91.67%, 65.71%, 52.94% for fungal detection across three groups. In summary, BALF mNGS demonstrates greater sensitivity in pathogen detection, albeit with reduced specificity in severe cases.

### Transcriptomic differences among mild CAP, SCAP (survivor), and SCAP (death) patients

PBMC from all mild CAP, SCAP (survivor), and SCAP (death) patients at admission were used for bulk RNA-sequencing to identify potential biomarkers of CAP severity and prognosis (Fig. 1). As illustrated in Fig. 4, DEGs were analyzed through pairwise comparisons of the CAP groups. In SCAP (death) versus mild CAP, 261 DEGs were up-regulated and 170 were down-regulated, while SCAP (survivor) had only two up-regulated DEGs. Comparisons between SCAP (death) and SCAP (survivor) revealed three up-regulated and one down-regulated DEG. FOLR3 (Folate Receptor 3) and ITGA7 (integrin subunit alpha 7) were up-regulated DEGs in both SCAP groups compared to mild CAP patients, with log2FoldChange values of 4.41 and 3.39 for FOLR3 and 3.02 and 2.89 for ITGA7, respectively (adjusted $P$-value < 0.001). The expression of OTOF (Otoferlin), SIGLEC1 (Sialic Acid Binding Ig Like Lectin 1), CXCL10 (C-X-C motif chemokine ligand 10), and MS4A4A (Membrane Spanning 4-Domains A4A) was significantly elevated in SCAP (death) compared to both mild CAP and SCAP (survivor). Cyclin A1(CCNA1) expression was higher in SCAP (death) compared to SCAP (survivor) and mild CAP but did not differ between SCAP (survivor) and mild CAP. In conclusion, FOLR3 and ITGA7 may serve as biomarkers of CAP severity, while OTOF, SIGLEC1, MS4A4A, and CXCL10 may be valuable for predicting both disease severity and poor outcomes in CAP.

### DEGs-based GO and KEGG pathway enrichment analysis

To identify key biological or molecular processes influencing CAP severity and prognosis, the identified DEGs were used to perform gene functional pathway enrichment analysis (Figs. 5 and 6). GO analysis was divided into three categories: biological process (BP), cellular component (CC), and molecular function (MF). The top up-regulated GO pathways in SCAP (survivor) patients compared to mild CAPs included folic acid transport, leukocyte migration, external side of plasma membrane, folic acid binding, and signaling receptor activity (Fig. 5A). Compared to mild CAP patients, the commonly up-regulated GO pathways in SCAP (death) primarily involved inflammatory response, extracellular region, collagen-containing extracellular matrix, extracellular space, extracellular matrix organization and RNA nuclease activity. The main down-regulated GO pathways included positive regulation of natural killer cell mediated cytotoxicity, adaptive immune response, T cell activation, immune response, external side of plasma membrane, and MHC class I protein complex binding (Fig. 5B, C). In addition, GO functional pathways, including cellular response to interleukin-17, endothelial cell activation, plasma membrane raft, CXCR3 chemokine receptor binding, and virion binding were up-regulated
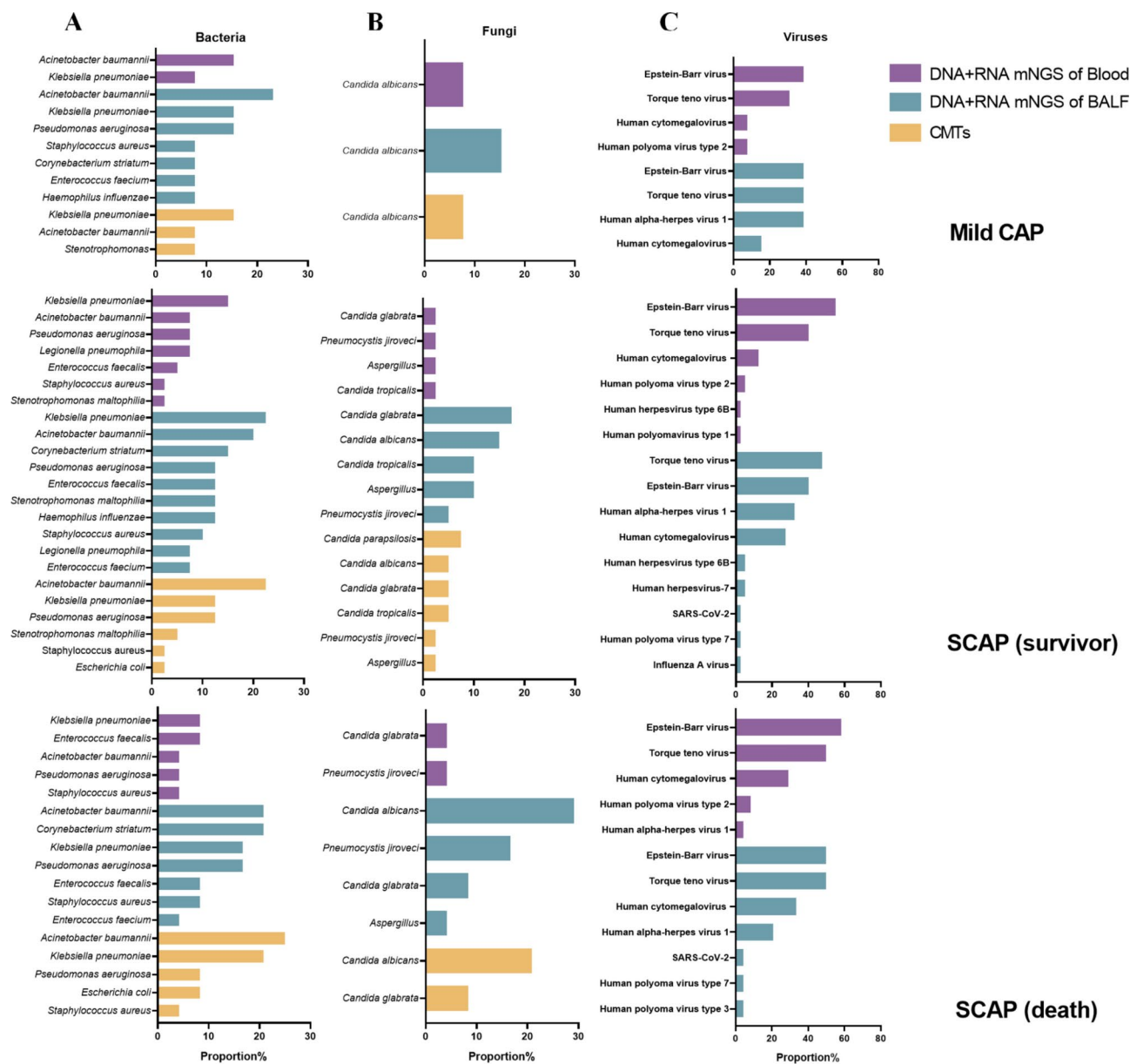
**Fig. 3.** Pathogen Spectrum in Mild CAP, SCAP (survivor), and SCAP (death) Patients Identified by Blood and BALF DNA + RNA mNGS Versus CMTs.

| Pathogens | SCAP (survivor) vs Mild CAP | SCAP (death) vs Mild CAP | SCAP (death) vs SCAP (survivor) |
|---|---|---|---|
| RNA + DNA mNGS of BALF | | | |
| Corynebacterium striatum | 15% vs 7.69% (P = 0.4990) | 20.83% vs 7.69% (P = 0.3005) | 20.83% vs 15% (P = 0.5493) |
| Candida albicans | 15% vs 15.38% (P = 0.9732) | 29.17% vs 15.38% (P = 0.3509) | 29.17% vs 15% (P = 0.1726) |
| Pneumocystis Jiroveci | 16.67% vs 0% (0.5649) | 16.67% vs 0% (P = 0.2756) | 16.67% vs 2.5% (P = 0.0409) |
| Human cytomegalovirus (HCMV) | 27.5% vs 15.38% (P = 0.3778) | 33.33% vs 15.38% (P = 0.2405) | 33.33% vs 27.5% (P = 0.6210) |
| RNA + DNA mNGS of blood | | | |
| Human cytomegalovirus (HCMV) | 12.5% vs 7.69% (0.6346) | 29.17% vs 7.69% (0.1298) | 29.17% vs 12.5% (0.0982) |
| Torque teno virus (TTVs) | 40% vs 30.77% (P = 0.5508) | 50% vs 30.77% (P = 0.2597) | 50% vs 40% (P = 0.4350) |
| Epstein-Barr virus (EBV) | 55% vs 38.46% (P = 0.3001) | 58.33% vs 38.46% (P = 0.2483) | 58.33% vs 55% (P = 0.7947) |

**Table 4.** Pathogen-Associated Risk Factors for Severe CAP and Prognosis: An RNA + DNA mNGS Analysis of BALF and Blood Samples.
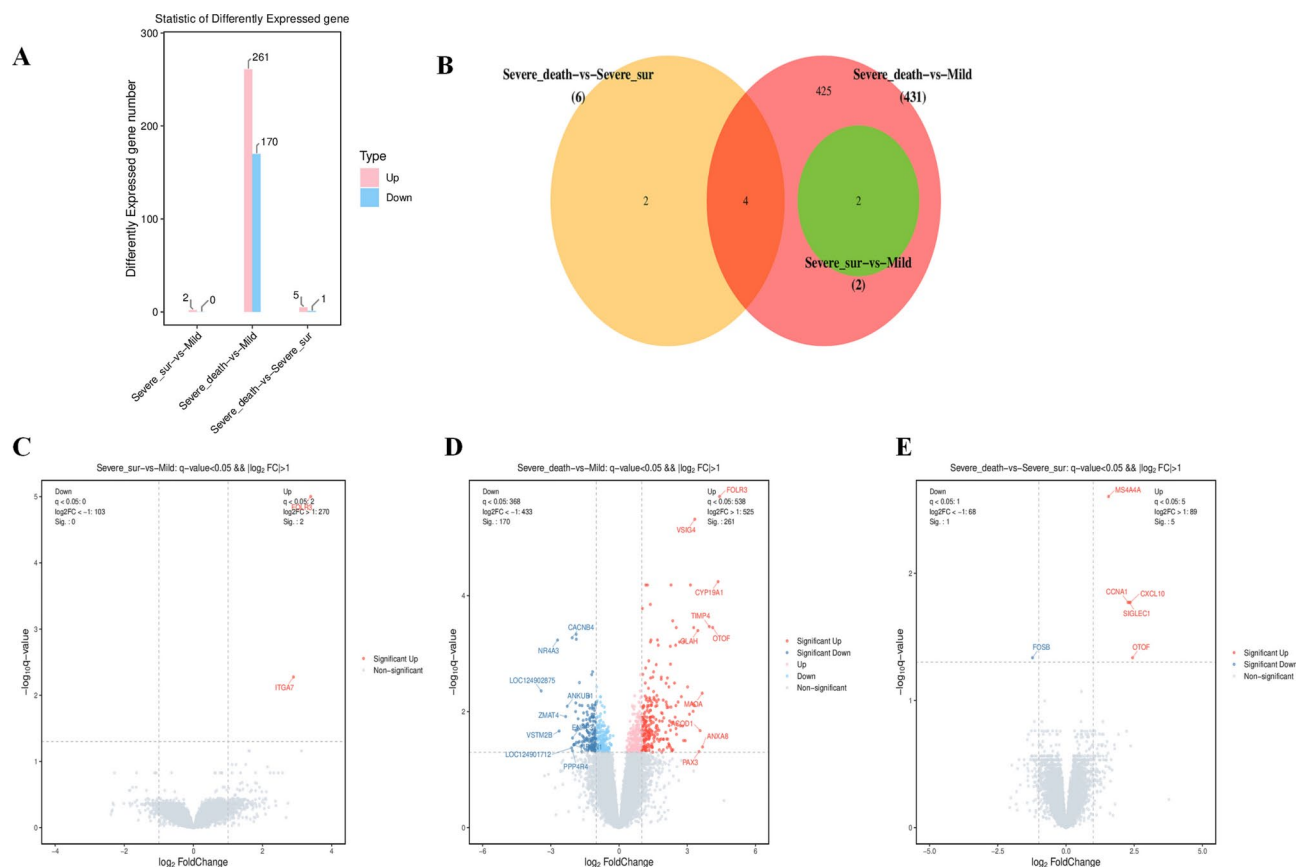
**Fig. 4**. Differentially Expressed Genes (DEGs) Analysis Among Mild CAP, SCAP (survivor), and SCAP (death) Patients. (**A**) Bar chart showing the number of down-regulated and up-regulated DEGs in comparisons of SCAP (Survivor) vs. Mild CAP, SCAP (Non-Survivor) vs. Mild CAP, and SCAP (Non-Survivor) vs. SCAP (Survivor). (**B**) Venn diagram illustrating the overlap in the number of DEGs across pairwise comparisons: SCAP (Survivor) vs. Mild CAP, SCAP (Non-Survivor) vs. Mild CAP, and SCAP (Non-Survivor) vs. SCAP (Survivor). (**C**) Volcano plot depicting the distribution of DEGs between SCAP (Survivor) and Mild CAP patients. (**D**) Volcano plot showing the distribution of DEGs between SCAP (Non-Survivor) and Mild CAP patients. (**E**) Volcano plot presenting the distribution of DEGs between SCAP (Non-Survivor) and SCAP (Survivor) patients.

in SCAP(death) compared to SCAP (survivor) patients (Fig. 5D), while the down-regulated pathway was the response to corticosterone.

The most impacted pathways associated with these common DEGs were also identified based on KEGG databases. KEGG pathways such as antifolate resistance, ECM-receptor interaction, and focal adhesion were identified as activated and played crucial roles in SCAP (survivor) compared to mild cases (Fig. 6A). Complement and coagulation cascades, nitrogen metabolism, and ECM-receptor interaction were the top up-regulated KEGG pathways enriched in SCAP (death) patients compared to mild cases, while the top down-regulated KEGG pathways included hematopoietic cell lineage, cell adhesion molecules, antigen processing and presentation, primary immunodeficiency, cytokine-cytokine receptor interaction, and T cell receptor signaling pathway (Fig. 6B, C). The up-regulated genes in SCAP (death) patients were mainly associated with Epstein-Barr virus infection, RIG-I-like receptor signaling pathway, and cytosolic DNA-sensing pathway when compared to mild patients (Fig. 6D). To further substantiate the conclusions of the pathway enrichment analysis, detailed statistical data, including *p*-values, adjusted *p*-values, gene ratios, and background ratios, have been compiled in the Appendix Tables 8 and 9.

### Gene set enrichment analysis (GSEA) of KEGG and GO pathways

GSEA KEGG pathway analysis, as shown in Appendix Table 10, revealed significant pathway enrichment in SCAP (death) and SCAP (survivor) groups compared to mild CAP. Immune-related pathways, including complement activation and neutrophil extracellular trap formation, were enriched in both groups, underscoring the crucial role of immune activation in severe pneumonia. In the SCAP (death) group, pathways related to energy metabolism, such as oxidative phosphorylation and lysosome function, were significantly upregulated, indicating increased cellular energy demand and stress response. In contrast, the SCAP (survivor) group showed downregulation of immune-related pathways, such as T cell receptor signaling and FoxO signaling, pointing to the importance of immune suppression and metabolic regulation for survival. The upregulation of
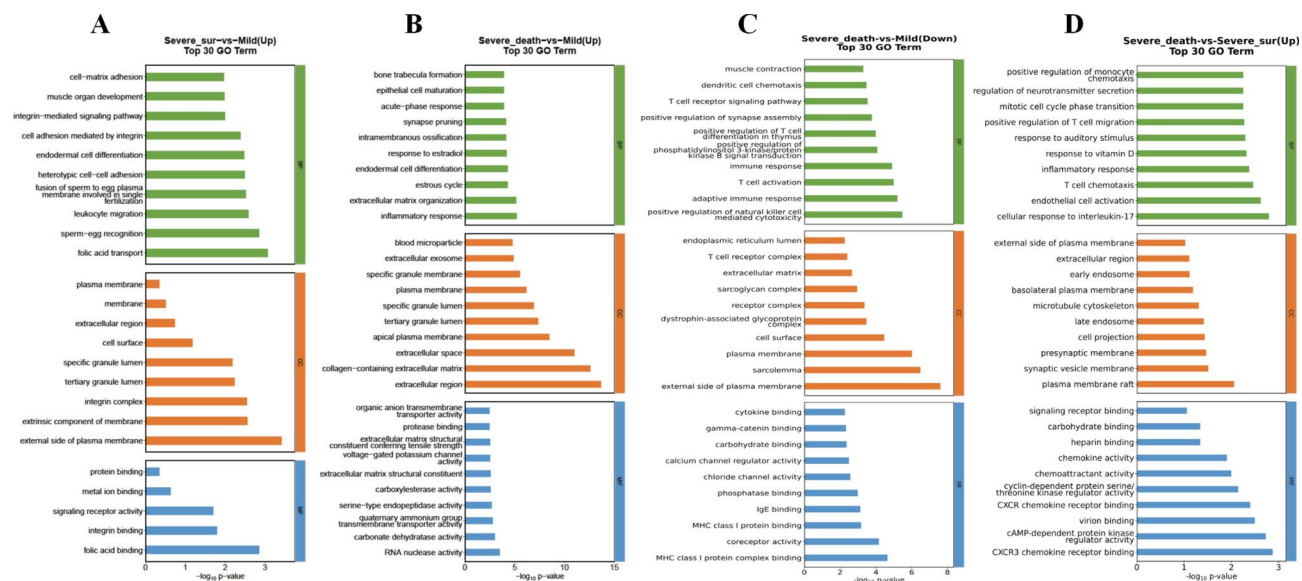
**Fig. 5.** GO Functional Pathway Enrichment Analysis in Mild CAP, SCAP (survivor), and SCAP (death) Patients (**A**) Top 30 upregulated GO functional pathways in SCAP (survivor) compared to Mild CAP patients. (**B**)Top 30 upregulated GO functional pathways in SCAP (death) compared to Mild CAP patients. (**C**)Top 30 downregulated GO functional pathways in SCAP (death) compared to Mild CAP patients. (**D**)Top 30 upregulated GO functional pathways in SCAP (death) compared to SCAP (survivor) patients.

oxidative phosphorylation, protein synthesis, and immune regulation pathways in SCAP (death) links metabolic dysregulation and excessive immune activation to mortality.

GSEA GO analysis of SCAP (death) vs. mild CAP revealed immune suppression and metabolic dysregulation as key features of severe pneumonia, including downregulation of T cell receptor signaling and T cell co-stimulation, and upregulation of proton-driven ATP synthesis, highlighting the role of immune decline and metabolic imbalance. The death group also showed significant upregulation in stress response pathways such as complement activation and protein folding, while downregulation of RNA transport and chromatin remodeling suggests immune and repair dysfunction as key factors in mortality. In SCAP (survivor) vs. mild CAP, the survivor group exhibited significant enrichment in acute phase response and cell cycle pathways, indicating the crucial role of immune regulation and cell repair for survival. Upregulation of low-density lipoprotein binding and carrier receptor activity supports energy supply for survival. Comparing SCAP (death) with SCAP (survivor), the death group showed significant upregulation of key pathways related to energy metabolism and immune response, such as proton-driven ATP synthesis, mitochondrial electron transport, and aerobic respiration, suggesting metabolic dysregulation as a direct cause of mortality. Excessive immune activation, particularly complement activation and telomere organization, points to immune overactivation accelerating death.

## Discussion

In this study, we characterized the pathogen spectra of mild CAP, SCAP (survivor) and SCAP (death) patients as identified by CMTs and DNA + RNA mNGS of BALF and blood, and compared their detection performance. The microbiome in BALF and blood, PBMC transcriptome profiles, along with clinical and laboratory characteristics of mild CAP and SCAP patients, were also analyzed to identify the potential biomarkers, clinical indicators or key functional pathways influencing disease severity and prognosis.

Overall, we found that incorporating BALF mNGS, as opposed to blood mNGS, into routine diagnosis workflows significantly enhanced the detection of bacteria, fungi and viruses. Timely and accurate identification of pathogens in CAP, particularly SCAP, is essential for determining the appropriate treatment plan and improving prognosis[9,10]. Although few previous study have compared the performance of blood and BALF mNGS in the same SCAP population, numerous studies have shown that BALF mNGS alone exhibits high diagnostic performance in LRTI patients and can provide clues for identifying rare pathogens such as *Mycobacterium abscessus*, *P. jiroveci*, *Orientia tsutsugamushi*, *Chlamydia psittaci*[9,25]. A meta-analysis reported that the pooled sensitivity and specificity of BALF mNGS for penumonia were 78% (95% confidence interval [CI] 67–87%; I2 = 92%) and 77% (95% CI 64–94%; I2 = 74%), respectively[26]. It was found that the sensitivity of BALF mNGS on viruses was much higher than that of CMTs (92.31% versus 7.69%, $P < 0.01$)[27]. Similarly, our study found that both blood DNA + RNA mNGS and BALF DNA + RNA mNGS performed excellently for DNA viruses but the detection rate of RNA virus was very low. The small genomes of most RNA viruses, their low abundance, the high degradation rate of viral RNA relative to host RNA, and inappropriate RNA library preparation methods may contribute to the low detection rate of RNA viruses by mNGS[28,29]. While previous studies have reported a positive detection rate of up to 92% (95% CI 78–100%) for pathogens using mNGS in BALF samples from severely or immunocompromised pulmonary-infected patients[26], our subgroup analyses demonstrated that the
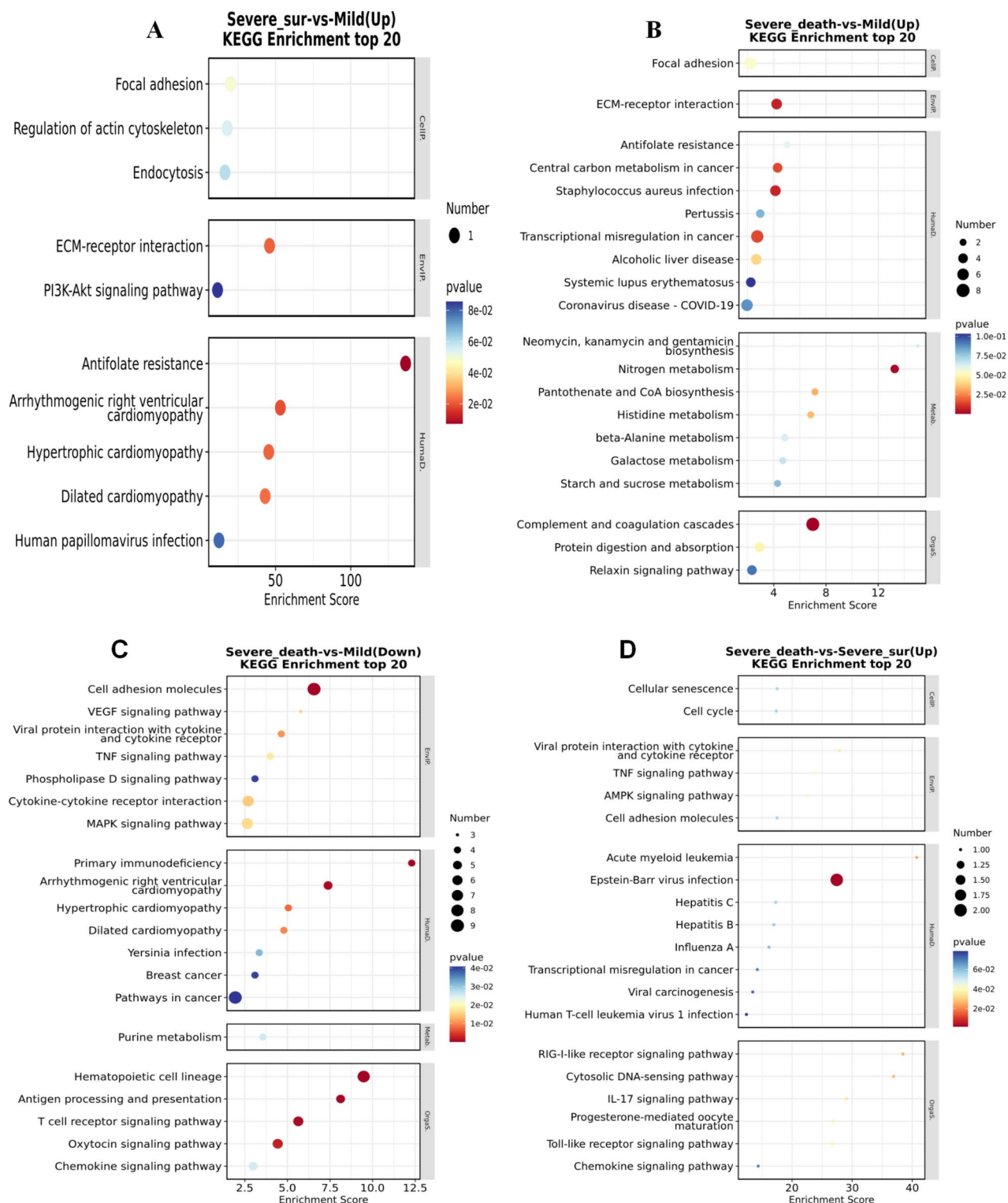
**Fig. 6**. KEGG Pathway Enrichment Analysis of PBMC Transcriptomes in Mild CAP, SCAP (survivor), and SCAP (death) Patients. (**A**) Top 20 upregulated KEGG functional pathways in SCAP (survivor) compared to Mild CAP patients. (**B**) Top 20 upregulated KEGG functional pathways in SCAP (death) compared to Mild CAP patients. (**C**) Top 20 downregulated KEGG functional pathways in SCAP (death) compared to Mild CAP patients. (**D**) Top 20 upregulated KEGG functional pathways in SCAP (death) compared to SCAP (survivor) patients.

diagnostic performance of mNGS in blood or BALF for bacterial pathogens was influenced by disease severity and declined among critical ill CAP patients, such as those in the SCAP (death) group. BALF or blood samples from critically ill patients are often characterized by significant infiltration of host inflammatory and tissue cells, leading to a high burden of host DNA[30]. This substantial presence of host DNA can markedly impede the accurate detection of bacterial DNA[31,32]. Consequently, the combination of CMTs and BALF mNGS is strongly recommended.

We identified the most common pathogens in SCAP as bacteria, including *A. baumannii*, *K. pneumoniae*, *Pseudomonas aeruginosa*, and *C. striatum*; fungi, including *C. albicans*, *P. jiroveci*, *C. glabrata*, *Candida tropicalis*, and *Aspergillus*; and viruses, including EBV, TTVs, HCMV and human polyomavirus type 2 (PyV-2). A national surveillance study in China reported that the most frequent pathogens associated with SCAP among older adults were influenza virus (10.94%) and *P. aeruginosa* (15.37%). In adults, SCAP was significantly associated with infection by *P. aeruginosa*, *K. pneumoniae*, or *S. pneumoniae*, or co-infection with *P. aeruginosa* and *K. pneumoniae*[33]. Our findings underscore the diverse pathogen landscape in SCAP and the critical need for targeted surveillance and management strategies, particularly for bacterial and viral co-infections in high-risk populations.

Contrary to previous research that suggested pathogen category or microbial diversity might not be linked to 30-day mortality in SCAP[14], we found that the infections by *P. jiroveci*, *C. striatum*, *C. albicans*, HCMV may indicate a worse prognosis in pneumonia patients. The poor prognosis of *P. jiroveci*-induced SCAP is primarily due to impaired host immunity, severe pulmonary pathology, delayed diagnosis, treatment challenges, and the presence of comorbidities and complications[34]. *C. striatum* detected by mNGS was once considered a contaminant because it is commonly found in the normal flora of the skin and oropharynx. However, recent studies have identified *C. striatum* as a pathogen in various infections including prosthetic joint infection, central line-associated bacteremia, endocarditis, and pleuropulmonary infections[18,35]. A retrospective study conducted in Korea from 2014 to 2019 indicated that *C. striatum* is emerging as a potential cause of severe pneumonia, particularly in immunocompromised patients[18]. To date, few studies have focused on *C. striatum* pneumonia, and the pathogenic mechanisms underlying SCAP caused by *C. striatum* remain to be explored. Severe pneumonia itself can lead to immune dysfunction, increasing susceptibility to opportunistic infections such as *C. albicans* and *P. jirovecii*, especially in critically ill patients[36,37]. The causality between severe pneumonia and the exacerbation by pathogens like *C. striatum* and *C. albicans* is challenging. In critically ill pneumonia patients, immune dysfunction is often a consequence of the disease itself, potentially increasing susceptibility to opportunistic infections. This highlights the complexity of distinguishing primary pathogens from secondary infections during pneumonia progression. Accelerated T-cell immunosenescence has been reported in HCMV-seropositive individuals[38]. Immune deficiency and severe infection can trigger the reactivation of latent HCMV infection, and the emergence of HCMV DNAemia in mild CAP patients is a rare event[39]. In general, the detection of HCMV DNAemia occurs predominantly in patients with PSI classes indicating more severe pneumonia[39]. Overall, our findings suggest that the detection of *C. striatum*, *C.albicans*, *P. jiroveci*, and HCMV by BALF or blood mNGS in SCAP patients generally indicates immune dysfunction and is associated with a worse prognosis.

According to RNA-seq analysis of PBMC, there were 431 DEGs in SCAP (death) compared to mild patients, however, few DEGs were found between SCAP (survivor) and mild CAP or between SCAP (death) and SCAP (survivor). Genes such as OTOF, MS4A4A, CXCL10, and SIGLEC1 serve as potential biomarkers for SCAP and its poorer outcomes. OTOF, involved in calcium ion signaling, may influence cellular responses to infection[40]. MS4A4A, known for its role in immune cell function modulation, could impact the immune response in severe CAP[41]. CXCL10 plays a critical role in recruiting immune cells to sites of inflammation, making it a reliable marker for predicting worse outcomes in SCAP[42]. SIGLEC1, expressed as CD169 on circulating monocytes, is involved in immune regulation and has been linked to disease severity in COVID-19, illustrating its potential as a biomarker for assessing severity and prognosis in SCAP[43,44]. So far, CXCL10 and SIGLEC1 have been more extensively studied and have established connections to severe inflammatory and infectious diseases, the roles of MS4A4A and OTOF in SCAP remain underexplored[40–42,44].

In SCAP (death) patients, GO pathways such as extracellular region, and collagen-containing extracellular matrix were up-regulated, while immune response, adaptive immune response, and dendritic cell chemotaxis pathway were down-regulated. Abnormal remodeling of the extracellular matrix (ECM) in the lung was associated with inflammatory response and fibrosis, which may contribute to severe COVID-19, tuberculosis, invasive pneumococcal disease (IPD)[45–47]. In addition, KEGG pathways such as complement and coagulation cascades, nitrogen metabolism, ECM-receptor interaction, and the EBV infection pathway were significantly up-regulated, while pathways involved in hematopoietic cell lineage, cell adhesion molecules, and antigen processing and presentation were down-regulated. Previously, a dysregulated complement system, coagulation cascade, and platelet function has been found to be potentially associated with multi-organ injury or failure in severe COVID-19 or pediatric SCAP[48–50]. Metabolic dysregulation, such as altered nitrogen metabolism, affects multiple organs. The versatile chemistry of nitrogen is important for pulmonary physiology and is involved in a range of pulmonary diseases, including asthma, cystic fibrosis, and adult respiratory distress syndrome[51,52]. EBV, a member of the herpesvirus family, is generally considered the cause of infectious mononucleosis (IM), lymphomas, and nasopharyngeal carcinomas[53]. EBV reactivation has been found to be common in sepsis patients and is associated with longer ICU stays (12.9 vs. 9.2 days) and increased organ failure (day 1 SOFA score 6.9 vs. 5.9)[54]. Similarly, we observed a higher proportion of EBV infections in SCAP compared to mild CAP, and the pathway associated with EBV infection was up-regulated in both SCAP survivors and non-survivors. Both GO and KEGG analyses suggest that a secondary immunosuppressive state and a weakened innate and adaptive immune system might be common features of SCAP (death), indicating a poorer outcome. In summary, RNA-

seq analysis of PBMCs provided molecular insights into the potential mechanisms of SCAP, identifying new biomarkers of disease severity and prognosis for early and precise intervention.

GSEA KEGG analysis revealed significant increase of immune-related pathways, such as complement activation and neutrophil extracellular trap formation, in the SCAP (death) group, that excessive immune activation may accelerate disease progression and contribute to mortality[55]. Furthermore, GSEA heightened activity in energy metabolism pathways, including oxidative phosphorylation and lysosomal function, in the SCAP (death) group, which reflects the increased cellular energy demands, consistent with the metabolic dysregulation observed in traditional analyses[56]. The interaction between metabolic dysregulation and immune overactivation likely causes cellular damage and exacerbates the disease. GSEA GO analysis further highlighted a reduction in T cell receptor signaling and metabolic pathways (e.g., proton-driven ATP synthesis) in the SCAP (death) group, while stress response pathways, including complement activation and protein folding, were notably increased, underlining the critical role of immune suppression and repair dysfunction in mortalityy[57,58]. In contrast, the SCAP (survivor) group showed enrichment in acute phase response and cell cycle pathways, emphasizing the importance of immune regulation and cellular repair for survival. By integrating GSEA with traditional GO/KEGG analysis, we gained a more comprehensive understanding of the roles of immune and metabolic dysregulation in severe pneumonia and identified potential therapeutic targets and prognostic markers for precision treatment.

Older age, a history of allergies, and cerebrovascular disease were clinical markers associated with a poorer outcome in SCAP patients. Laboratory indicators such as WBC count, neutrophil count, ALT, AST, and CRP initially increased and then decreased with the progression of CAP, warranting attention when assessing disease severity and predicting prognosis. An acute and critical condition should be considered in SCAP patients who despite being identified as SCAP by scoring systems such as the PSI score, CURB-65 score, or SOFA score, present with unusually low levels of WBC count, neutrophil count, ALT, AST, and CRP. This phenomenon is akin to the "silent chest" observed in critical asthma patients, which may reflect a state of multiple organ failure and myelosuppression. Additionally, laboratory indicators such as FDP, D-dimer, PRO-BNP, CK-MB, Myoglobin, serum IgE, and β-D-1, 3-glucan (fungus) gradually increased with disease severity and may be positively correlated with poorer outcomes. We hypothesized that coagulation abnormalities, cardiac insufficiency, fungal co-infection, and allergic status might have a mutual causality with SCAP severity and play a critical role in the progression of CAP.

Our study has several strengths. First, this research was a prospective multicenter study, which enhances its reliability and applicability. Second, enrolled patients received both blood and BALF mNGS simultaneously, enabling a pairwise comparison of the performance of blood mNGS, BALF mNGS, CMTs among SCAP patients. Third, our study linked detailed clinical features, laboratory indicators, and multi-omics bio-signatures, including DNA + RNA mNGS and PBMC RNA-seq analysis, to the differential severity and prognosis of SCAP patients. Our study also has some limitations. First, the relatively small size of the cohort may limit the generalizability of our results. Larger, multicenter studies will be needed to validate these findings across broader clinical settings. Moreover, the detection rate of RNA virus was low; therefore, the current library preparation and sequencing processes for DNA + RNA mNGS should be optimized. Besides, dynamic changes in multi-omics bio-signatures during SCAP progression were not assessed due to the difficulty of specimen acquisition. Finally, our sample collection occurred during COVID-19 management, excluding COVID-19 patients. With shifting global policies, changes in pathogen spectra and immune responses could affect our findings. Future studies should account for these factors.

## Conclusions
This study investigated the performance of BALF and blood DNA + RNA mNGS, CMTs in pathogen detection among patients with mild CAP and SCAP. Furthermore, the clinical features, laboratory indicators, and multi-omics bio-signatures of the microbiome and PBMC transcriptome were analyzed to identify biomarkers for disease severity and prognosis. The performance of BALF mNGS was superior to that of blood mNGS, and the detection efficiency of mNGS for bacteria was influenced by disease severity. Therefore, we highly recommended the combination of BALF mNGS and CMTs for SCAP diagnosis. Microbiome biomarkers such as *P. Jiroveci, C. striatum, C. albicans,* HCMV, EBV, as well as transcriptome bio-signatures such as OTOF, MS4A4A, CXCL10, and SIGLEC1, were associated with poorer outcomes in SCAP. GSEA and traditional GO/KEGG analyses revealed key immune and metabolic dysregulation pathways in SCAP (death) patients, including upregulated pathways related to complement and coagulation cascades, oxidative phosphorylation, nitrogen metabolism, and EBV infection, while adaptive immune response, hematopoietic cell lineage, and antigen processing pathways were downregulated. Overall, microbiomic and transcriptomic analyses provide a molecular pathological basis for identifying biomarkers associated with CAP severity and prognosis which is beneficial for individualized treatment.

## Data availability
All data are available in the main text or the Supplementary Materials. Correspondence and requests for raw sequencing data should be addressed to Yanan Li and Jieming Qu. The sequencing data submitted in the Sequence Read Archive (SRA) submission number is PRJNA1177187.

# References

1. Gupta, A. B. et al. Inappropriate diagnosis of pneumonia among hospitalized adults. *JAMA Intern. Med.* **184**, 548–556 (2024).
2. Ramirez, J. A. et al. Adults hospitalized with pneumonia in the United States: Incidence, epidemiology, and mortality. *Clin. Infect Dis.* **65**, 1806–1812 (2017).
3. Jain, S. et al. Community-acquired pneumonia requiring hospitalization among US adults. *N. Engl. J. Med.* **373**, 415–427 (2015).
4. Martin-Loeches, I. et al. ERS/ESICM/ESCMID/ALAT guidelines for the management of severe community-acquired pneumonia. *Intensive Care Med.* **49**, 615–632 (2023).
5. Qu, J. et al. Aetiology of severe community acquired pneumonia in adults identified by combined detection methods: A multi-centre prospective study in China. *Emerging Microb. Infect.* **11**, 556–566 (2022).
6. Olson, G. & Davis, A. M. Diagnosis and treatment of adults with community-acquired pneumonia. *JAMA* **323**, 885–886 (2020).
7. Hu, X. et al. The diagnostic value of bronchoalveolar lavage fluid metagenomic next-generation sequencing in critically ill patients with respiratory tract infections. *Microbiol. Spectrum* **12**, e0045824 (2024).
8. Wu, X. et al. Effect of metagenomic next-generation sequencing on clinical outcomes of patients with severe community-acquired pneumonia in ICU: A multicenter randomized controlled trial. *Chest* **167**, 362–373 (2025).
9. Diao, Z., Han, D., Zhang, R. & Li, J. Metagenomics next-generation sequencing tests take the stage in the diagnosis of lower respiratory tract infections. *J. Adv. Res.* **38**, 201–212 (2022).
10. Sun, T. et al. Metagenomic next-generation sequencing for pathogenic diagnosis and antibiotic management of severe community-acquired pneumonia in immunocompromised adults. *Front. Cell Infect. Microbiol.* **11**, 661589 (2021).
11. Wang, C. et al. Clinical application of bronchoalveolar lavage fluid metagenomics next-generation sequencing in cancer patients with severe pneumonia. *Respir Res.* **25**, 68 (2024).
12. Wei, Y. et al. Clinical evaluation of metagenomic next-generation sequencing for the detection of pathogens in BALF in severe community acquired pneumonia. *Ital. J. Pediatr.* **49**, 25 (2023).
13. Wu, X. et al. Etiology of severe community-acquired pneumonia in adults based on metagenomic next-generation sequencing: A prospective multicenter study. *Infect Dis. Ther.* **9**, 1003–1015 (2020).
14. Zhao, J. et al. A multicenter prospective study of comprehensive metagenomic and transcriptomic signatures for predicting outcomes of patients with severe community-acquired pneumonia. *EBioMedicine* **96**, 104790 (2023).
15. Hoffmann, M. et al. Blood transcriptomics analysis offers insights into variant-specific immune response to SARS-CoV-2. *Sci. Rep.* **14**, 2808 (2024).
16. Wang, K. et al. Sequential multi-omics analysis identifies clinical phenotypes and predictive biomarkers for long COVID. *Cell Rep Med.* **4**, 101254 (2023).
17. Bernardes, J. P. et al. Longitudinal multi-omics analyses identify responses of megakaryocytes, erythroid cells, and plasmablasts as hallmarks of severe COVID-19. *Immunity* **53**, 1296-1314.e1299 (2020).
18. Lee, Y. W. et al. Severe pneumonia caused by corynebacterium striatum in adults, Seoul, South Korea, 2014–2019. *Emerg. Infect Dis.* **28**, 2147–2154 (2022).
19. Mendes, A. R., Costa, A., Ferreira, H. & Ferreira, C. Enterococcus faecalis-associated lung abscess in a male adolescent—a case report. *BMC Pediatr.* **20**, 98 (2020).
20. Solis, M. et al. Combining predictive markers for severe COVID-19: Torquetenovirus DNA load and SARS-CoV-2 RNAemia. *J. Clin. Virol.* **148**, 105120 (2022).
21. Moss, B. J. & Musher, D. M. Candida species in community-acquired pneumonia in patients with chronic aspiration. *Pneumonia* **13**, 12 (2021).
22. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: Biological systems database as a model of the real world. *Nucl. Acids Res.* **53**, D672–D677 (2025).
23. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).
24. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* **28**, 27–30 (2000).
25. Lin, T. et al. Microbiological diagnostic performance of metagenomic next-generation sequencing compared with conventional culture for patients with community-acquired pneumonia. *Front. Cell. Infect. Microbiol.* **13**, 1136588 (2023).
26. Chen, S., Kang, Y., Li, D. & Li, Z. Diagnostic performance of metagenomic next-generation sequencing for the detection of pathogens in bronchoalveolar lavage fluid in patients with pulmonary infections: Systematic review and meta-analysis. *Int. J. Infect. Dis IJID Off. Publ. Int. Soc. Infect. Dis.* **122**, 867–873 (2022).
27. Jin, X. et al. Improving suspected pulmonary infection diagnosis by bronchoalveolar lavage fluid metagenomic next-generation sequencing: A multicenter retrospective study. *Microbiol. Spectr.* **10**, e0247321 (2022).
28. He, Y. et al. Enhanced DNA and RNA pathogen detection via metagenomic sequencing in patients with pneumonia. *J. Transl. Med.* **20**, 195 (2022).
29. Marston, D. A. et al. Next generation sequencing of viral RNA genomes. *BMC Genomics* **14**, 444 (2013).
30. Long, M. E., Mallampalli, R. K. & Horowitz, J. C. Pathogenesis of pneumonia and acute lung injury. *Clin. Sci.* **136**, 747–769 (2022).
31. Hogan, C. A. et al. Clinical impact of metagenomic next-generation sequencing of plasma cell-free DNA for the diagnosis of infectious diseases: A multicenter retrospective cohort study. *Clin. Infect. Dis.* **72**, 239–245 (2021).
32. Li, N. et al. High-throughput metagenomics for identification of pathogens in the clinical settings. *Small Methods* **5**, 2000792 (2021).
33. Liu, Y.-N. et al. Infection and co-infection patterns of community-acquired pneumonia in patients of different ages in China from 2009 to 2020: A national surveillance study. *Lancet Microbe.* **4**, e330–e339 (2023).
34. Ghembaza, A., Vautier, M., Cacoub, P., Pourcher, V. & Saadoun, D. Risk factors and prevention of *P. jirovecii* pneumonia in patients with autoimmune and inflammatory diseases. *Chest* **158**, 2323–2332 (2020).
35. Söderquist, B., Henningsson, T. & Stegger, M. Corynebacterium striatum prosthetic joint infection successfully treated with long-term dalbavancin. *Microorganisms* **11**, 550 (2023).
36. Li, Y.-N. et al. Inhibition of myeloid-derived suppressor cell arginase-1 production enhances T-cell-based immunotherapy against *Cryptococcus neoformans* infection. *Nat. Commun.* **13**, 4074 (2022).
37. Cano-Gamez, E. et al. An immune dysfunction score for stratification of patients with acute infection based on whole-blood gene expression. *Sci. Transl. Med.* **14**, 4433 (2022).
38. Álvarez-Heredia, P. et al. Accelerated T-cell immunosenescence in cytomegalovirus-seropositive individuals after severe acute respiratory syndrome coronavirus 2 infection. *J. Infect. Dis.* **228**, 576–585 (2023).
39. Saletinger, R., Poljak, M. & Strle, F. Presence of human cytomegalovirus DNA in blood of patients with community-acquired pneumonia. *Clin. Microbiol. Infect.* **21**, 97 (2015).
40. Vincent, P. F. Y. et al. Different CaV1.3 channel isoforms control distinct components of the synaptic vesicle cycle in auditory inner hair cells. *J Neurosci* **37**, 2960–2975 (2017).
41. Mattiola, I., Mantovani, A. & Locati, M. The tetraspan MS4A family in homeostasis, immunity, and disease. *Trends Immunol.* **42**, 764–781 (2021).
42. MacLean, A. J. et al. Secondary influenza challenge triggers resident memory B cell migration and rapid relocation to boost antibody secretion at infected sites. *Immunity* **55**, 718 (2022).
43. Doehn, J.-M. et al. CD169/SIGLEC1 is expressed on circulating monocytes in COVID-19 and expression levels are associated with disease severity. *Infection* **49**, 757–762 (2021).

44. Zheng, Q. et al. Siglec1 suppresses antiviral innate immune response by inducing TBK1 degradation via the ubiquitin ligase TRIM27. *Cell Res.* **25**, 1121–1136 (2015).
45. Murphy, S. L. et al. Circulating markers of extracellular matrix remodelling in severe COVID-19 patients. *J. Intern Med.* **294**, 784–797 (2023).
46. Hartog, C. M. et al. Pulmonary matrix metalloproteinase excess in hospital-acquired pneumonia. *Am. J. Respir. Crit. Care Med.* **167**, 593–598 (2003).
47. Trappetti, C., Ogunniyi, A. D., Oggioni, M. R. & Paton, J. C. Extracellular matrix formation enhances the ability of Streptococcus pneumoniae to cause invasive disease. *PLoS ONE* **6**, e19844 (2011).
48. Wang, Y. et al. Serum-integrated omics reveal the host response landscape for severe pediatric community-acquired pneumonia. *Crit. Care* **27**, 79 (2023).
49. Moraes, E. C. D. S. et al. Proteomic profile of procoagulant extracellular vesicles reflects complement system activation and platelet hyperreactivity of patients with severe COVID-19. *Front. Cell. Infect. Microbiol.* **12**, 926352 (2022).
50. Huber, S. et al. Systemic inflammation and complement activation parameters predict clinical outcome of severe SARS-CoV-2 infections. *Viruses* **13**(12), 2376. https://doi.org/10.3390/v13122376 (2021).
51. Marozkina, N. V. & Gaston, B. Nitrogen chemistry and lung physiology. *Annu. Rev. Physiol.* **77**, 431–452 (2015).
52. Fan, J. et al. GlnR activated transcription of nitrogen metabolic pathway genes facilitates biofilm formation by mycobacterium abscessus. *Int. J. Antimicrob. Agents* **63**, 107025 (2024).
53. Niazi, M. R. et al. Epstein-Barr virus (EBV) induced pneumonitis in an immunocompetent adult: A case report. *Respir. Med. Case Rep.* **31**, 101262 (2020).
54. Goh, C. et al. Epstein-Barr virus reactivation in sepsis due to community-acquired pneumonia is associated with increased morbidity and an immunosuppressed host transcriptomic endotype. *Sci. Rep.* **10**, 9838 (2020).
55. Grégoire, M. et al. Impaired efferocytosis and neutrophil extracellular trap clearance by macrophages in ARDS. *Eur. Respir. J.* **52**, 1702590 (2018).
56. Ning, P. et al. Metabolic profiles in community-acquired pneumonia: Developing assessment tools for disease severity. *Crit Care* **22**, 130 (2018).
57. Reizine, F. et al. Beneficial effects of citrulline enteral administration on sepsis-induced T cell mitochondrial dysfunction. *Proc. Natl. Acad. Sci. USA* **119**, e2115139119 (2022).
58. Baral, P. et al. Nociceptor sensory neurons suppress neutrophil and γδ T cell responses in bacterial lung infections and lethal pneumonia. *Nat. Med.* **24**, 417–426 (2018).

## Author contributions

Yannan Li and Jieming Qu conceived and designed the study. Wanmei Song and Qing-yuan Yang directed its implementation including the data analysis and writing of the paper. Wanmei Song and Hui Lv analyzed the data; Hui Lv, Yingqi Lv and Yanshan Jiang contributed materials/analytic tools; Wanmei Song and Qingyuan Yang wrote and revised the manuscript. All authors revised it critically for important intellectual content, gave final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved.

## Declarations

### Ethics approval

This study was approved by the Ruijin Hospital Ethics Committee, Shanghai Jiao Tong University School of Medicine (Ethics Approval No.: Ruijin Hospital Ethics Committee 2021-73); Shanghai General Hospital, Shanghai Jiao Tong University (Ethics Approval No.: 2021KY041); Zhongshan Hospital, Fudan University (Ethics Approval No.: B2021-183R); Xinhua Hospital, Shanghai Jiao Tong University School of Medicine (Ethics Approval No.: XHEC-C-2021-038-1); Shanghai Chest Hospital, Shanghai Jiao Tong University (Ethics Approval No.: 2021-130); Shanghai Public Health Clinical Center (Ethics Approval No.: 2021-S035-02); Renji Hospital, Shanghai Jiao Tong University School of Medicine (Ethics Approval No.: KY2021-103-B); and Shanghai Sixth People's Hospital, Shanghai Jiao Tong University (Ethics Approval No.: IS2147). The study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Written informed consent was obtained from all participants prior to inclusion in the study.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-00812-x.

**Correspondence** and requests for materials should be addressed to Y.L.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.