BMJ Open Respiratory Research

Clinical characteristics, diagnosis, outcomes and lung microbiome analysis of invasive pulmonary aspergillosis in the community-acquired pneumonia patients

Zhi Ao,[1] Huan Xu,[2] Mengqi Li,[1] Huifang Liu,[2] Min Deng,[1] Yuliang Liu[1]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjresp-2022-001358).

Check for updates

¹Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China
²Department of Scientific Affairs, Vision Medicals for Infection Diseases, Guangzhou, China

Correspondence to
Dr Yuliang Liu;
37428643@qq.com

ABSTRACT

Background Invasive pulmonary aspergillosis (IPA) remains underestimated in patients with community-acquired pneumonia (CAP). This study aims to describe clinical features and outcomes of IPA in CAP patients, assess diagnostic performance of metagenomic next-generation sequencing (mNGS) for IPA and analyse lung microbiome via mNGS data.

Methods This retrospective cohort study included CAP patients from 22 April 2019 to 30 September 2021. Clinical and microbiological data were analysed. Diagnostic performance of mNGS was compared with traditional detection methods. The lung microbiome detected by mNGS was characterised and its association with clinical features was evaluated.

Main results IPA was diagnosed in 26 (23.4%) of 111 CAP patients. Patients with IPA displayed depressed immunity, higher hospital mortality (30.8% vs 11.8%) and intensive care unit mortality (42.1% vs 17.5%) compared with patients without IPA. The galactomannan (GM) antigen test had the highest sensitivity (57.7%) in detecting the *Aspergillus* spp, followed by mNGS (42.3%), culture (30.8%) and smear (7.7%). The mNGS, culture and smear had 100% specificity, while GM test had 92.9% specificity. The microbial structure of IPA significantly differed from non-IPA patients (p<0.001; Wilcoxon test). Nineteen different species were significantly correlated with clinical outcomes and laboratory biomarkers, particularly for *Streptococcus salivarius*, *Prevotella timonensis* and *Human betaherpesvirus 5*.

Conclusions Our results reveal that patients with *Aspergillus* infection tend to have a higher early mortality rate. The mNGS may be suggested as a complement to routine microbiological test in diagnosis of patients at risk of *Aspergillus* infection. The lung microbiota is associated with inflammatory, immune and metabolic conditions of IPA, and thus influences clinical outcomes.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Detection methods including galactomannan antigen, smear and culture have been widely used in the diagnosis of invasive pulmonary aspergillosis (IPA). Bacterial 16S ribosomal RNA gene sequencing recently has been used in the analysis of lung microbiota in IPA. However, no study to date has reported the application of metagenomic next-generation sequencing (mNGS) in the diagnosis of IPA and evaluation of lung microbiota in IPA.

WHAT THIS STUDY ADDS

⇒ We showed that community-acquired pneumonia (CAP) patients with IPA displayed suppressed immunity, high hospital mortality and intensive care unit mortality. The mNGS exhibited high sensitivity and specificity in diagnosis of IPA. Moreover, dysbiosis of lung microbiome was related to early mortality, ventilation, oxygen supply, inflammatory and immunity condition and metabolism.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The results of this study are important for a better understanding IPA in the context of CAP. The mNGS may be suggested as a complement to routine microbiological test in diagnosis of patients at risk of *Aspergillus* infection to increase detection rate. The lung microbiota provided a new perspective for the pathogenesis of aspergillosis and may represent a novel therapeutic target for the prevention and treatment of *Aspergillus* infection.

INTRODUCTION

Invasive pulmonary aspergillosis (IPA) is an important cause of morbidity and mortality, typically affecting immunocompromised hosts, such as patients with HIV infection as well as patients receiving corticosteroids therapy, transplant recipients or haematologic cancer.[1] Numerous cases of misdiagnosis or missed diagnosis and only diagnosed post-mortem due to lack of medical resource and new risk factors are underestimated.[2] Clinical diagnosis of IPA comprises host factors, clinical features and mycological evidence.[3] Problems arising from the inaccessibility of tissue

samples, low sensitivity of culture, poor applicability of traditional host factors and absence of typical clinical and radiological findings lead to delayed and missed diagnosis in IPA patients.[4] The recently published American Thoracic Society clinical practice guidelines for the diagnosis of fungal infection highlight the culture-independent diagnostic tests, with strong recommendations given to the detection of *Aspergillus* DNA in bronchoalveolar lavage fluid (BALF) by PCR.[5]

In the past decade, development in molecular methods for the quantification and sequencing of pathogenic DNA has greatly facilitated the detection of various microorganisms. Metagenomic next-generation sequencing (mNGS) has been now widely applied to the diagnosis of infectious diseases in clinical settings, especially with specific or unidentified or mixed pathogens in a non-targeted approach.[6] For example, combining mNGS with conventional detection methods could increase the detection rate for mycobacterium tuberculosis complex[7] and facilitate the detection of mixed pathogens in immunocompromised severe community-acquired pneumonia (CAP) patients.[8] To date, the diagnostic performance of mNGS for detection of lung *Aspergillosis* infection has not been clear.

Recently, variations in composition and diversity of lung microbiome were described across numerous respiratory diseases, providing new sights for precise diagnosis. Patients with bacteriologically confirmed tuberculosis exhibited a considerable increase in bacterial species compared with negative tuberculosis patients.[9] The imbalance of pulmonary community driven by enrichment with potential pathogen is correlated with outcomes in intubated/mechanically ventilated patients.[10] Hérivaux *et al* discovered patients with IPA exhibited a differential abundance of specific bacterial taxa.[11]

With newer diagnostic techniques available, it is feasible to evaluate clinical characteristics and lung microbiome of IPA from a new perspective for a comprehensive understanding of the disease. This study was designed to describe clinical characteristics and outcomes of IPA in CAP patients, assess diagnostic performance of mNGS for IPA and reveal the relationship between clinical characteristics and lung microbiome in IPA.

METHODS
Study design and participants
This was a retrospective cohort study that consisted of hospitalised patients (≥18 years of age) diagnosed as CAP at The First Affiliated Hospital of Chongqing Medical University between 22 April 2019 and 30 September 2021. Patients enrolled in this study should meet the following inclusive criteria: (1) diagnosed with CAP meeting the 2007 IDSA/ATS guidelines[12], (2) undergoing bronchoalveolar lavage, (3) no antibiotic exposure in the last 2 weeks before sample collection and (4) having galactomannan (GM) test, smear, culture and BALF mNGS results of BALF. Cases with probable IPA require the presence of at least one host factor, a clinical manifestation and mycologic evidence in BALF, according to the 2019 criteria from the European Organization for Research and Treatment of Cancer/Mycoses Study Group (EORTIC/MSG).[3] In this study, a modification in diagnostic criteria was the addition of mNGS as one of mycologic testing methods including GM test, smear and culture. Cases without probable IPA included patients with no evidence of clinical or radiological evidence or the presence of *Aspergillus* spp in the BALF. All the enrolled patients were followed up by telephone to inquire about survival or date of death.

Data collection and microbiological testing
Clinical information was extracted from the electronic medical records using a standardised data collection form, which included demographic information, clinical data, laboratory findings, CT of the chest, antifungal treatments and outcomes. All the data were reviewed by two investigators independently to verify data accuracy. The 7-category ordinal scale[13] was modified based on the study design, which consisted of mutually exclusive categories as follows: category 7, death; category 6, requiring invasive ventilation and additional organ support (eg, continuous renal replacement therapy or extracorporeal membrane oxygenation); category 5, requiring invasive mechanical ventilation; category 4, requiring non-invasive ventilation or high-flow oxygen; category 3, requiring supplemental oxygen by mask or nasal prongs; category 2, hospitalised, no oxygen; category 1, discharged. Improvement in clinical status was defined as a decline of 2 categories assessed by the 7-level ordinal scale at fixed time points (day 1, 7, 14, 21 and 28).

Testing for GM antigen in serum and GM in BALF is generally performed in all enrolled patients. A result was considered positive when optical index values were ≥1.0 in serum or ≥1.0 in BALF or ≥0.7 in serum and ≥0.8 in BALF. BALF specimens were subjected to mNGS, smear by microscopy and ordinary fungal culture.

Metagenomics next-generation sequencing
A 1.5–3 mL of BALF was collected from each patient according to standard procedures.[14] A 1.5 mL microcentrifuge tube containing 0.6 mL BALF sample, enzyme and 1 g of 0.5 mm glass beads was agitated by a vortex mixer, and a 0.3 mL sample was separated into a 1.5 mL microcentrifuge tube and DNA was extracted using the TIANamp Micro DNA Kit (DP316, Tiangen Biotech) following the manufacturer's manual.[15] DNA libraries were constructed through DNA fragmentation, end repair, adapter ligation and PCR amplification. Then, libraries with confirmed quality were sequenced by the illumina NextSeq 550 platform with single reads of 75 bp (Illumina, San Diego, California, USA).[16]

The internal control, named unique molecular spiked-in (UMSI), was added to the sample before the DNA extraction. The sequence of UMSI varied in

different samples. Each mNGS assay run included an external negative control that ran in parallel with clinical samples. During analysis, the contamination between samples could be found if the UMSI sequence was the same or the reads of some pathogens in the external control were very high.

High-quality sequencing data were generated by removing low-quality reads, followed by computational subtraction of human host sequences mapped to the human reference genome (hg19) using Burrows-Wheeler alignment.[17] The remaining data by removal of low-complexity reads were classified by simultaneously aligning to four microbial genome databases (bacteria, fungi, viruses and parasites), which were downloaded from the National Center for Biotechnology Information (ftp://ftp.ncbi.nlm.nih.gov/genomes/). Interpretation of mNGS results can be seen in online supplemental methods.

Patient and public involvement

Patients and the public were not involved in the design and conduct of this study, choice of outcome measures, nor recruitment. The results of this study will not be disseminated to the participant and linked communities.

Statistical analyses

Continuous variables were expressed as median (IQR) and compared with the Mann-Whitney U test. Categorical variables were expressed as number (%) and compared by χ^2 test or Fisher's exact test. The 2×2 contingency tables were established to determine sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). A two-sided α of less than 0.05 was considered statistically significant. Statistical analyses were done using the IBM SPSS Statistics, V.23.0, unless otherwise indicated.

The high-quality reads were first mapped to the human reference genome (hg19) and then non-human reads were aligned to four microbial genome databases (bacteria, fungi, viruses and parasites). After removing the background microbial which appeared in negative controls (online supplemental table 1), Bayesian re-estimation of abundance was used to estimate the relative abundance of species (online supplemental table 2). The alpha diversity index for each sample was calculated based on Shannon and Simpson indexes, and beta diversity was calculated based on weighted UniFrac distance and Bray-Curtis distance. P values for alpha and beta analysis were calculated using Wilcoxon rank-sum test. Linear discriminant analysis (LDA) effect size was used to find significantly different species between groups, with thresholds of log10 LDA Score≥2 and p value≤0.05. The correlation between the representative microbiota and clinical data was evaluated by Spearman.

RESULTS

Patient recruitment and clinical characteristics

A total of 123 patients with 128 BALF specimens were reviewed and 111 patients were enrolled in this study.
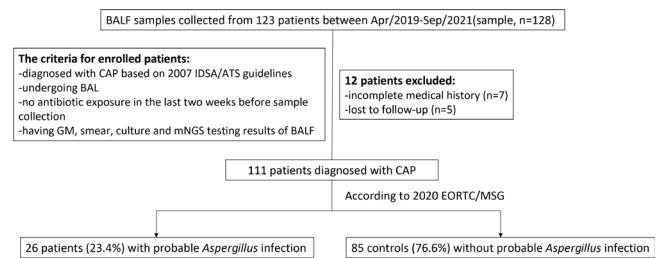

Figure 1 Flowchart of selected cases. BAL, bronchoalveolar lavage; BALF, bronchoalveolar lavage fluid; CAP, community-acquired pneumonia; EORTC/MSG, the European Organization for Research and Treatment of Cancer and the Mycoses Study Group; GM, galactomannan; IDSA/ATS, Infectious Diseases Society of America/American Thoracic Society; mNGS, metagenomic next-generation sequencing.

Twenty-six patients were diagnosed with probable IPA (named IPA) and 85 controls without IPA (named control) (figure 1). Comparisons of demographic and clinical characteristics between IPA group and control group were summarised in table 1. The study population had a median age of 60 (IQR, 53–73) years and was 66.7% were men. Overall, 33 patients (29.7%) had a history of smoking. Twenty-two patients (19.8%) had chronic obstructive pulmonary disease. Fifteen patients (13.5%) had received corticosteroids and 36 patients (32.4%) were considered to be immunocompromised (online supplemental table 3). Patients with probable IPA had a significantly higher rate with known risk factors for aspergillosis than those without IPA, including treatment with corticosteroids longer than 28 days (38.5% vs 5.9%, p=0.000) and immunocompromised status (53.8% vs 25.9%, p=0.008). No significant differences were found in sequential organ failure assessment scoring and confusion, urea, respiratory rate, blood pressure, age 65 scoring of patients between two groups on admission.

Microbiological, CT images and laboratory findings

Aspergillus fumigatus isolates were obtained for 6 of 26 (23.1%) probable IPA cases, followed by *Aspergillus flavus* for 3 of 26 (11.5%) and *Aspergillus ustus* for 1 of 26 (3.8%) (table 2). Bronchial wall thickening was more often shown on CT of the chest in the IPA (44.0% vs 12.8%, p=0.000) (table 2). IPA patients have significantly lower level of total lymphocyte count (774.1 vs 1088.9, p=0.009) and lymphocyte subpopulation, including $CD3^+$ T lymphocyte count (553.0 vs 782.6, p=0.016), $CD4^+$ T lymphocyte count (287.9 vs 409.5, p=0.031) and $CD4^- CD8^-$ T lymphocyte count (25.9 vs 45.0, p=0.007) than control patients (online supplemental table 4), indicating the compromised immune function of IPA patients.

Comparison of diagnostic performance

To compare the diagnostic performance of BALF mNGS and conventional tests for distinguishing cases with probable IPA from cases without probable IPA, the results were

Table 1 Demographics and clinical characteristics of the study cohort

	Suspected CAP cohort (n=111)	With probable *Aspergillus* infection (n=26; 23.4%)	Without probable *Aspergillus* infection (n=85; 76.6%)	P value
Baseline factors				
Age, year, median (IQR)	60 (53–73)	64 (53–74)	59 (50–72)	0.58
Sex, Male, n (%)	74 (66.7)	14 (53.8)	60 (70.6)	0.113
Pregnancy/recent delivery, n (%)	1 (0.9)	0	1 (1.2)	1.000
Smoking, n (%)	33 (29.7)	5 (19.2)	28 (32.6)	0.181
Dyslipidaemia, n (%)	7 (6.3)	1 (3.8)	6 (7.1)	0.898
Arterial hypertension, n (%)	31 (27.9)	6 (23.1)	25 (29.4)	0.619
Diabetes mellitus, n (%)	25 (22.5)	5 (19.2)	20 (23.5)	0.529
Chronic liver failure, n (%)	2 (1.8)	0	2 (2.4)	1.000
Chronic heart failure, n (%)	11 (9.9)	3 (11.5)	8 (9.4)	1.000
Chronic kidney failure, n (%)	4 (3.6)	1 (3.8)	3 (3.5)	1.000
Known risk factors for aspergillosis				
COPD, n (%)	22 (19.8)	3 (11.5)	19 (22.3)	0.353
Haematologic malignancy, n (%)	3 (2.7)	1 (3.8)	2 (2.3)	1.000
Autoimmune disease, n (%)	19 (17.1)	7 (26.9)	12 (14.1)	0.129
AIDS, n (%)	1 (0.9)	1 (3.8)	0	0.234
Active solid tumour, n (%)	15 (13.5)	6 (23.1)	9 (10.6)	0.103
Solid-organ transplant, n (%)	2 (1.8)	1 (3.8)	1 (1.1)	0.958
Neutropenia, n (%)	2 (1.8)	2 (7.7)	0	0.053
Corticosteroids 28 days, n (%)	15 (13.5)	10 (38.5)	5 (5.9)	0.000
Other immunosuppressive therapy, n (%)	16 (14.4)	6 (23.1)	10 (11.8)	0.151
Immunocompromised*, n (%)	36 (32.4)	14 (53.8)	22 (25.9)	0.008
Severity scoring				
Median SOFA Score at day 1 (IQR)	3 (1–5)	3 (2–5)	3 (1–4)	0.212
Median CURB-65 Score at day 1 (IQR)	1 (0–2)	1 (0–2)	1 (0–2)	0.308

Data are given as n (%) or median (IQR), unless otherwise specified. When no denominator is specified, the denominator is the number of participants in the corresponding cell in the first row.
P values were calculated by Fisher's exact probability t-test for categorical variables or Mann-Whitney U test for continuous variables.
*Immunocompromised: patients meeting the 2019 European Organization for Research and Treatment of Cancer host factors.
CAP, community-acquired pneumonia; COPD, chronic obstructive pulmonary disease; CURB-65, confusion, urea, respiratory rate, blood pressure, age 65; SOFA, sequential organ failure assessment.

showed in figure 2. GM test (57.7%) had the highest sensitivity in detecting the *Aspergillus* spp, followed by mNGS (42.3%), culture (30.8%) and smear (7.7%). Comparing with smear, mNGS had the significantly higher sensitivity 42.3% vs 7.7%, p=0.01). The mNGS, culture and smear had 100% specificity, while GM test had 92.9% specificity, with significant difference between mNGS (or culture or smear) and GM test (p=0.029). The PPV of mNGS, culture and smear for identifying *Aspergillus* spp was 100%, while GM test was 71.4%. The NPV ranked as GM test (87.8%), mNGS (85.0%), culture (82.5%) and smear (78.0%), with no significant difference among these methods. We have analysed differences in sensitivity/specificity/PPV/ NPV of GM on serum versus BALF, using a BAL fluid

GM index of ≥1.0 or serum GM index of ≥1.0 as positive, according to 2019 EORTIC/MSG. It was significantly higher in sensitivity but lower in specificity of GM BALF than that of GM serum (57.7% vs 26.9%, p=0.048, 92.9% vs 100%, 0.029, respectively), which was shown in online supplemental table 5. There were no differences in PPV and NPV.

Clinical outcomes

There were no significant differences in 28-day, 60-day and 90-day mortality between the two groups (table 3). Hospital mortality (30.8% vs 11.8%, p=0.021) and intensive care unit (ICU) mortality rates (42.1% vs 17.5%,

Table 2 Microbiological findings and CT images of the study cohort

	Suspected CAP cohort (n=111)	With probable *Aspergillus* infection (n=26; 23.4%)	Without probable *Aspergillus* infection (n=85; 76.6%)	P value
Aspergillus spp				
Aspergillus fumigatus	6 (5.5)	6 (23.1)	0	0.000
Aspergillus flavus	3 (1.8)	3 (11.5)	0	0.012
Aspergillus ustus	1 (0.09)	1 (3.8)	0	0.234
Thoracic CT findings				
Nodules	24 (21.6)	7 (28.0)	17 (19.8)	0.453
Micronodules	11 (9.9)	2 (8.0)	9 (10.5)	0.954
Bronchial wall thickening	22 (19.8)	12 (44.0)	10 (12.8)	0.000
Consolidation	79 (71.2)	18 (72.0)	61 (70.9)	0.803
Cavity or air-crescent sign	17 (15.3)	7 (28.0)	10 (11.6)	0.060
Ground glass opacity	64 (57.7)	17 (64.0)	47 (55.8)	0.362

Data are given as n (%), unless otherwise specified. When no denominator is specified, the denominator is the number of participants in the corresponding cell in the first row. P values were calculated by Fisher's exact probability t-test for categorical variables.
CAP, community-acquired pneumonia.

p=0.029) were significantly higher in the IPA group than in the control group. The distribution of patients falling into each category of the 7-category scale showed statistical difference between the two groups by day 28 (p=0.017); however, significance with statistical testing was not reached in the 7-category ordinal scale by days 1, 7 and 14 (figure 3). Multiple antifungal drugs were administered for 23.1% of patients with probable IPA, compared with 5.9% of patients without probable IPA (p=0.01). The proportion of patients receiving voriconazole was remarkably higher in cases with probable IPA (53.9% vs 8.2%, p=0.000). There was no significant difference in the use of amphotericin B and caspofungin between the two groups (table 3).

Figure 2 Diagnostic performance of bronchoalveolar lavage fluid mNGS and conventional testing for differentiating probable invasive pulmonary aspergillosis (IPA) from non-probable IPA. Gm, galactomannan; mNGS, metagenomic next-generation sequencing; neg, negative; NPV, negative predictive value; pos, positive; PPV, positive predictive value.

Lung microbiome analysis

To compare the overall composition and diversity of the lung microbial signature in patients with IPA and the control, we analysed BALF specimens collected from 109 CAP patients, including 24 cases (missing 2 patients due to loss of mNGS data) diagnosed with probable IPA (named IPA) and 85 cases without probable IPA (named control). We assessed α diversity of the lung communities using the Shannon and Simpson diversity index. There was a decrease in α diversity both in Shannon and Simpson diversity index in IPA patients as compared with non-IPA patients in spite of significant difference (figure 4A,B). Using the weighted UniFrac metric and Bray-Curtis metric, we observed that β diversity of IPA patients significantly differed from non-IPA patients (p<0.001; Wilcoxon test), suggesting that lung community structure of patients diagnosed with probable IPA differed substantially from those without probable IPA (figure 4C,D). Based on the average relative abundance, we plotted the top 10 phyla, genus and species among 2 groups. In the overall cohort, *Firmicutes*, *Proteobacteria*, *Actinobacteria*, *Bacteroidetes* and *Ascomycota* were the most abundant phyla, whereas *Prevotella*, *Streptococcus*, *Acinetobacter*, and *Pneumocystis* were the most common genera (figure 5A,B). The relative abundance of the top 10 in species ranked as *Pneumocystis jirovecii*, *Acinetobacter baumannii*, *Lautropia mirabilis*, *Streptococcus oralis*, *Corynebacterium striatum*, *Human betaherpesvirus 5*, *Rothia mucilaginosa*, *Staphylococcus aureus*, *Prevotella melaninogenica*, *Nocardia Farcinica*, with only *H. betaherpesvirus 5* was the significantly different between two groups (figure 5C). Focusing on the differential species in the IPA compared with their controls, we found 21 species with LDA Score≥2 and p<0.05, which *H. betaherpesvirus 5*, *A. fumigatus*, *Aspergillus niger*, *Citrobacter braakii*, *Bacillus thermoamylovorans*, *Helcococcus kunzii*, *Lactobacillus*

Table 3 Clinical outcomes of the study cohort

	Suspected CAP cohort (n=111)	With probable *Aspergillus* infection (n=26; 23.4%)	Without probable *Aspergillus* infection (n=85; 76.6%)	P value
Mortality				
Mortality (28 days)	34 (30.6)	12 (46.2)	22 (25.9)	0.05
Mortality (60 days)	1 (0.9)	0	1 (1.2)	1.000
Mortality (90 days)	4 (3.6)	0	4 (4.7)	0.571
Hospital mortality	18 (16.2)	8 (30.8)	10 (11.8)	0.021
ICU mortality (76)	18/76 (23.7)	8/19 (42.1)	10/57 (17.5)	0.029
Prognosis				
Improvement in clinical status, median (IQR), days	14 (7–19)	13.5 (0–23)	14 (7–18)	0.848
Ventilator-free days within 28 days, median (IQR)	26 (18–28)	24 (12–28)	26 (19–28)	0.290
Supplemental oxygen independence days within 28 days, median (IQR), days	14 (8–21)	12.5 (4.5–23.5)	15 (10–21)	0.457
Duration of ICU stay, median (IQR), days	7 (0–14)	12 (0–16)	7 (0–12)	0.087
Duration of hospital stay, median (IQR), days	16 (12–22)	18.5 (13–27)	16 (10–21)	0.058
Treatment				
Multiple antifungal drugs*	11 (10.8)	6 (23.1)	5 (5.9)	0.010
Main antifungal therapy†				
Voriconazole	21 (18.9)	14 (53.9)	7 (8.2)	0.000
Amphotericin B	3 (2.7)	1 (3.8)	2 (2.4)	1.000
Caspofungin	18 (16.2)	5 (19.2)	13 (15.3)	0.634

Data are n (%) or median (IQR), unless otherwise specified. When no denominator is specified, the denominator is the number of participants in the corresponding cell in the first row. Parenthetical values in the stub column represent the number of patients for whom data were available.
*Multiple antifungal drugs: used in combination or consecutively.
†Main antifungal therapy: first antifungal drug administered for ≥5 consecutive days following invasive pulmonary aspergillosis diagnosis.
CAP, community-acquired pneumonia; ICU, intensive care unit.

delbrueckii, Burkholderia dolosa, Marinobacter hydrocarbono-clasticus, Riemerella anatipestifer, Corynebacterium halotolerans and *Lactobacillus plantarum* were significantly abundant in the cases diagnosed with probable IPA, and *Streptococcus salivarius, Citrobacter freundii, Paraburkholderia fungorum, Dolosigranulum pigrum, Prevotella timonensis, Sphingobium yanoikuyae, Serratia marcescens* and *Corynebacterium oculi* were enriched in cases without probable IPA (figure 5D, online supplemental figure 1).

To further investigate the correlation between clinical data and the lung microbiota of IPA, we performed Spearman's rank-based correlation test on the 21 significantly different species with clinical outcome and laboratory findings (figure 6A). We observed 19 different species were significantly correlated with 7 clinical outcomes and 22 laboratory biomarkers, particularly *S. salivarius, P. timonensis* and *H. betaherpesvirus 5. S. salivarius* was primarily positively correlated with eight laboratory biomarkers, including red blood cell, total

lymphocytes, T lymphocytes, T helper cells, cytotoxic T lymphocytes, B lymphocytes, serum potassium and albumin, and negatively relative with aspartate aminotransferase, direct bilirubin and D-dimer. Whereas, *S. salivarius* was mainly negatively associated with 11 clinical outcomes, including ICU mortality, duration of ICU stay, hospital mortality, mortality at day 28 and time to supplemental oxygen independence within 28 days, and positively correlated with duration of hospital stay mortality at day 28, ventilator-free days within 28 days and time to improvement in clinical status. Similar with *S. salivarius, P. timonensis* was primarily positively relative six laboratory biomarkers, including total lymphocytes, T lymphocytes, T helper cells, CD4⁻CD8⁻ T cells, B lymphocytes and serum potassium, and negatively correlated with procalcitonin. Equally, *P. timonensis* was mainly negatively associated with duration of ICU stay and hospital mortality, but positively with ventilator-free days within 28 days. Opposite to *S. salivarius* and *P. timonensis, H.*
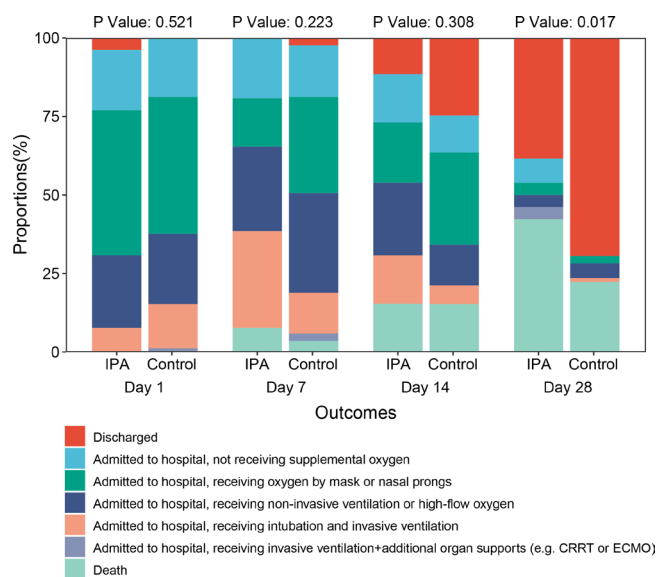
Figure 3 The clinical outcomes of IPA and control at days 1, 7, 14 and 28 according to the 7-category ordinal scale end point. P values were calculated using the Wilcoxon rank sum test. CRRT, continuous renal replacement therapy; ECMO, extracorporeal membrane oxygenation; IPA, invasive pulmonary aspergillosis.
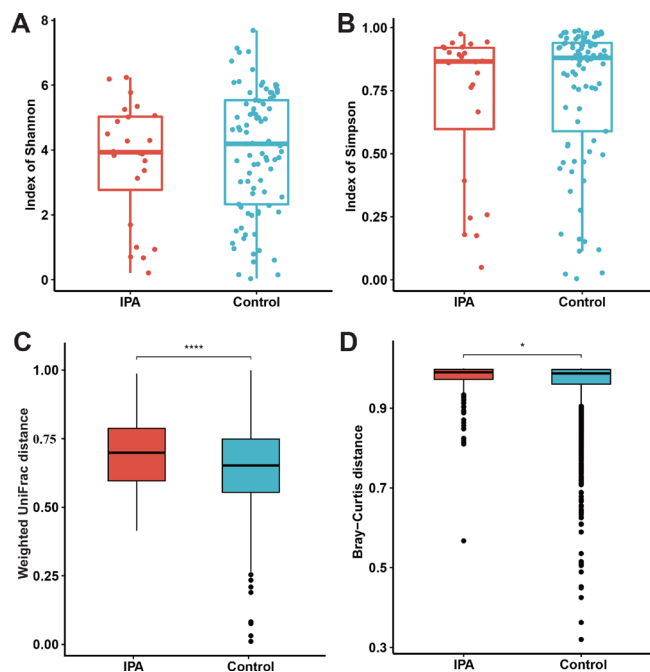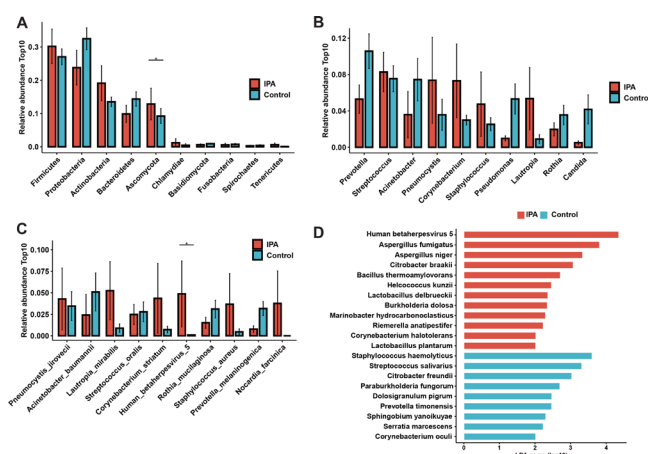

Figure 5 The microbial composition of invasive pulmonary aspergillosis (IPA) and control. The top 10 phylum (A), genus (B) and species (C) in relative abundance of IPA and control, and the significantly differential species (D) was plotted. P values for the top phylum, genus and species were calculated using the Wilcoxon rank sum test. The significantly differential species were calculated using LDA effect size (D), with thresholds of log10 LDA score≥2 and p value≤0.05. LDA, linear discriminant analysis. *P value<0.05; **p value<0.01; ***p value<0.001; ****p value<0.0001.

Figure 4 The alpha and beta diversity of invasive pulmonary aspergillosis (IPA) and control. Alpha and beta diversity of IPA (n=24) and control subjects (n=85). Shannon (A) and Simpson (B) indexes were used to calculate the alpha diversity community within each individual microbiota sample. Beta diversity between the IPA and control was depicted by weighted UniFrac distance (C) and Bray-Curtis distance (D). Data are presented as box plot overlaid by a dot plot with a line at the median. P values were calculated using the Wilcoxon rank sum test. *p value<0.05; ****p value<0.0001.

betaherpesvirus 5 was negatively associated with six laboratory biomarkers, including total lymphocyte, T lymphocytes, T helper cells, CD4$^+$CD8$^+$ T cells, B lymphocytes and serum potassium, and positively relative with blood urea nitrogen. Meanwhile, *H. betaherpesvirus 5* was mainly negatively correlated with duration of ICU stay and time to supplemental oxygen independence within 28 days. These observations suggested the dysbiosis of microbial communities in the lung is associated with the patient's pathophysiological conditions and thus can influence the clinical outcomes.

To further investigate the clinical impact of 21 significantly different species, we made the survival analysis by plotting the cumulative probability curve. As shown in figure 6B,C, detection of *A. fumigatus* and *H. betaherpesvirus 5* was significantly predictive of worse ICU outcomes (p=0.045, 0.032). While in figure 6D,E, we found that detection of *S. salivarius* and *P. timonensis* was extremely significantly predictive of improved ICU outcomes (p=0.0031, 0.0054). We thus concluded that in the lung microbiota of CAP, ICU outcomes may be predicted by community composition, specifically *A. fumigatus*, *H. betaherpesvirus 5*, *S. salivarius* and *P. timonensis*.

DISCUSSION

In this retrospective cohort study, we examined clinical characteristics of IPA. Interestingly, we noted that patients categorised as having probable *Aspergillus* infection had a significantly higher hospital mortality (30.8%) and ICU mortality (42.1%) compared with the control. The mortality at day 28, though not statistically significant,

Figure 6 The association between significantly differential species and the clinical data. The heat map shows the Spearman correlation between microbial species with clinical data (A). Cumulative incidences of intensive care unit (ICU) discharge alive rate among patients with *Human betaherpesvirus 5* (B) and *Streptococcus salivarius* (C) infection. Red values indicate species were positive correlated with clinical data, while blue ones indicate the species were negative correlated with clinical data. Significant associations (adjusted p<05) are indicated by asterisk. ALT, alanine aminotransferase; APTT, activated partial thromboplastin time; AST, aspartate aminotransferase; Blood_Urea, blood urea nitrogen; B_LY, B lymphocytes; CD3⁺_T, T lymphocytes; CD4⁺_T, T helper cells; CD8⁺_T, cytotoxic T lymphocytes; CD4⁺CD8⁺_T, CD4⁺CD8⁺ T lymphocytes; CD4⁻CD8⁻_T, CD4⁻CD8⁻ T lymphocytes; CRP, C reaction protein; DBIL, direct bilirubin; DHS, duration of hospital stay mortality at day 28; DICUS, duration of ICU stay; Hb, haemoglobin; HM, hospital mortality; ICS, time to improvement in clinical status; ICUM, ICU mortality; INR, international normalised ratio; LDH, lactate dehydrogenase; M28, mortality at day 28; Neu, neutrophil count; NK, nature killer cells; PCT, procalcitonin; RBC, red blood cell; Serum_ K⁺, serum potassium; Serum_ Cl⁻, serum chlorinum; Total LY, total lymphocytes; TSOI, time to supplemental oxygen independence within 28 days; VFD, ventilator-free days within 28 days; WBC, whole blood cell.

was numerically higher in IPA patients than the control. These data suggested that patients with *Aspergillus* infection tend to have a higher early mortality rate. Previous

study reported an estimated 45% of ICU mortality and 26% of hospital mortality[18] Among invasive tracheobronchial aspergillosis, a rare entity of IPA, 30-day and 90-day mortality rates were both as high as 90%.[19] Loughlin *et al* reported that ICU mortality in patients with probable *Aspergillus* infection was not significantly higher than those without *Aspergillus* infection,[4] which possibly due to generally critical status in ventilator-associated pneumonia patient cohort in the ICU setting. In this study, the overall condition of the enrolled patients with CAP was relatively mild, which can be remarkably aggravated by *Aspergillus* infection and potentially life-threatening.

Cornillet *et al* have suggested that IPA patients' non-specific clinical features and the diagnostic difficulties particularly when patients without neutropenia easily lead to suboptimal management and delayed initiation of treatment, resulting in high mortality and morbidity.[20] Given the devastating mortality rates of patients with IPA, timely and accurate diagnosis is highlighted and urgently needed for effective antifungal treatment to decrease the morbidity and mortality rate. In the past decade, GM detection has been widely used in the diagnosis of IPA.[21] However, GM testing for serum samples has limited sensitivity in diagnose IPA in non-neutropenic hosts.[22] Instead, detection of GM in BAL fluid samples of patients at risk of IPA has an excellent diagnostic accuracy,[23 24] because GM is released earlier and at higher concentrations in the BAL fluid in comparison with serum.[25] Recently, mNGS has been playing an increasingly important role in the diagnosis of fungal diseases. Ding *et al* demonstrated that the sensitivity of mNGS was significantly higher than culture in detecting *Candida albicans*.[9] In this study, BALF was collected directly from the lower respiratory tract, which avoided contamination from the oral microbiota. *Aspergillus* in BALF were detected by mNGS, culture, smear and GM test. The results showed that mNGS had comparable specificity and PPV to culture in *Aspergillus* detection. Although the sensitivity of mNGS lower than GM test, mNGS showed more sensitive than smear and more specific than serum and BALF GM test. These results indicated that mNGS acts as a useful non-culture based test for detecting *Aspergillus*, suggesting as a complement to routine microbiological test in diagnosis of patients at risk of IPA.

Our study discovered that CAP patients with *Aspergillus* were more likely to coinfected with bacterial or (and) viral infection due to the immunocompromised status and corticosteroid treatment. Patients with influenza who were immunocompromised were reported to have incidence of IPA was as high as 32%, whereas in the non-immunocompromised influenza case group, incidence was 14%.[18] With the emergence of the COVID-19 pandemic and corticosteroids use in patients with severe COVID-19,[26] COVID-19-associated pulmonary aspergillosis (CAPA) has been reported in a number of case reports.[27–29] In a meta-analysis, the incidence and mortality of CAPA in the ICU were estimated to be 10.2% and 54.9%, respectively.[30]

Due to the vast microbial reads in the mNGS data that can be used to analyse microbial signature across numerous diseases, our study attempted to explore the difference in structure and composition of microbiome between IPA patients and non-IPA CAP patients by analysing the mNGS data. Our present study comprised the lung microbiota of 24 IPA patients and 85 non-IPA CAP patients. The relationship of different microbiota and clinical data for all 109 CAP patients were investigated. For alpha diversity, there were no significant differences in the Shannon index and Simpson index of microbiota in IPA patients and non-IPA CAP patients, demonstrating that the with-in sample diversity of communities were similar between IPA patients and non-IPA patients. For beta diversity, our results displayed a dramatical difference in the weighted UniFrac distance ($p=2.1\times10^{-6}$) and Bray-Curtis distance ($p=0.033$) between IPA and non-IPA CAP patients, indicating a potential effect of *Aspergillus* spp on lung microbiota structure in patients diagnosed with IPA. Similar studies have showed significant differences in beta diversity of microbiota between IPA and controls.[11] This may be that the causative agents in the controls were complex, which can be any CAP pathogens (such as bacteria, virus, fungi, parasite and even atypical pathogens) except *Aspergillus*. Thus the lung microbial heterogeneity in the control was higher than IPA. Meanwhile, these studies demonstrated that *Firmicutes*, *Proteobacteria*, *Actinobacteria* and *Bacteroidetes* were the most abundant phyla in the IPA cohort.[11] However, beyond the bacteria, our study found that the relative abundance of *Ascomycota* was also high in the level of phylum, for the metagenomics can harvest almost all the nuclei acid of microbiota, including bacteria, fungi, virus and parasite.

The significantly differential species, particularly *S. salivarius*, *P. timonensis* and *H. betaherpesvirus 5*, were strongly related with clinical data. *S. salivarius* and *P. timonensis* were positively related with laboratory biomarkers such as T lymphocytes and negatively related with clinical outcomes such as hospital mortality. Whereas, *H. betaherpesvirus 5* was negatively related with laboratory biomarkers such as T lymphocytes and positively related with clinical outcomes such as hospital mortality. Previous study demonstrated that *S. salivarius* is a probiotic bacterium that can inhibit pneumococcal adherence to pharyngeal epithelial cells.[31] Also, another study proven that *S. salivarius* isolates had a superior ability to inhibit the growth of *Haemophilus influenzae*, *Moraxella catarrhalis*, *Streptococcus pneumoniae*, *Streptococcus pyogenes* and *Staphylococcus aureus*.[32] All these studies coincided our study that *S. salivarius* was positively correlated with the better clinical outcome of the patients. The possible mechanism of *S. salivarius* positively correlated with the improved outcome of the patients may be that they can inhibit the potential pathogens growth and activate the host immune system. For *P. timonensis*, previous study showed that *P. timonensi*s could dramatically induce the expression of membrane-associated mucins in the three-dimensional endometrial epithelial cell model and formed the biofilm in the epithelial surface, which in turn influence the adhesion of other bacteria.[33] The reason for the positive correlation with the better clinical outcome of the patients may be the colonisation resistance of *P. timonensis* in the lung and the activation of the host immune system. In clinical settings, *H. betaherpesvirus 5* can build life-long latent infection in the healthy people. However, once immunocompromised, *H. betaherpesvirus 5* can be reactivated, for example, solid organ transplant recipients and HIV-positive patients.[34] Thus, the reactivation of the *H. betaherpesvirus 5* indicated the weakened immune system and the poor clinical outcomes of IPA patients.

Our study has several limitations. First, the retrospective study was designed to analyse collected data and limited to relatively small sample size, which could predispose it to introducing bias from incomplete information. Second, this was the observation study. Though we found some interesting correlation between lung microbiota and clinical data, causal relationship needs to be resolved by prospective cohort study or mechanism experiment, such as animal experiment, which will be the next step of our study. Otherwise, the study identified patients with probable *Aspergillus* infection strictly based on new EORTIC/MSG 2019 diagnostic criteria so that a few cases could not satisfy the criteria for a probable IPA diagnosis due to absence of classic host factors.

REFERENCES

1 Taccone FS, Van den Abeele A-M, Bulpa P, *et al*. Epidemiology of invasive aspergillosis in critically ill patients: clinical presentation, underlying conditions, and outcomes. *Crit Care* 2015;19:7.
2 Danion F, Rouzaud C, Duréault A, *et al*. Why are so many cases of invasive aspergillosis missed? *Medical Mycology* 2019;57:S94–103.
3 Donnelly JP, Chen SC, Kauffman CA, *et al*. Revision and update of the consensus definitions of invasive fungal disease from the European organization for research and treatment of cancer and the mycoses Study Group education and research Consortium. *Clin Infect Dis* 2020;71:1367–76.
4 Loughlin L, Hellyer TP, White PL, *et al*. Pulmonary aspergillosis in patients with suspected ventilator-associated pneumonia in UK ICUs. *Am J Respir Crit Care Med* 2020;202:1125–32.
5 Hage CA, Carmona EM, Evans SE, *et al*. Summary for clinicians: microbiological laboratory testing in the diagnosis of fungal infections in pulmonary and critical care practice. *Annals ATS* 2019;16:1473–7.
6 Mitchell SL, Simner PJ. Next-generation sequencing in clinical microbiology: are we there yet? *Clin Lab Med* 2019;39:405–18.
7 Liu X, Chen Y, Ouyang H, *et al*. Tuberculosis diagnosis by metagenomic next-generation sequencing on bronchoalveolar lavage fluid: a cross-sectional analysis. *Int J Infect Dis* 2021;104:50–7.
8 Sun T, Wu X, Cai Y, *et al*. Metagenomic next-generation sequencing for pathogenic diagnosis and antibiotic management of severe community-acquired pneumonia in immunocompromised adults. *Front Cell Infect Microbiol* 2021;11:661589.
9 Ding L, Liu Y, Wu X, *et al*. Pathogen metagenomics reveals distinct lung microbiota signatures between bacteriologically confirmed and negative tuberculosis patients. *Front Cell Infect Microbiol* 2021;11:708827.
10 Dickson RP, Schultz MJ, van der Poll T, *et al*. Lung microbiota predict clinical outcomes in critically ill patients. *Am J Respir Crit Care Med* 2020;201:555–63.
11 Hérivaux A, Willis JR, Mercier T, *et al*. Lung microbiota predict invasive pulmonary aspergillosis and its outcome in immunocompromised patients. *Thorax* 2022;77:283–91.
12 Mandell LA, Wunderink RG, Anzueto A, *et al*. Infectious diseases society of america/american thoracic society consensus guidelines on the management of community-acquired pneumonia in adults. *Clin Infect Dis* 2007;44 Suppl 2:S27–72.
13 Wang Y, Fan G, Horby P, *et al*. Comparative outcomes of adults hospitalized with seasonal influenza A or B virus infection: application of the 7-category ordinal scale. *Open Forum Infect Dis* 2019;6:ofz053.
14 Chen X, Ding S, Lei C, *et al*. Blood and bronchoalveolar lavage fluid metagenomic next-generation sequencing in pneumonia. *Can J Infect Dis Med Microbiol* 2020;2020:6839103.
15 Long Y, Zhang Y, Gong Y, *et al*. Diagnosis of sepsis with cell-free DNA by next-generation sequencing technology in ICU patients. *Arch Med Res* 2016;47:365–71.
16 Jeon YJ, Zhou Y, Li Y, *et al*. The feasibility study of non-invasive fetal trisomy 18 and 21 detection with semiconductor sequencing platform. *PLoS ONE* 2014;9:e110240.
17 Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 2009;25:1754–60.
18 Schauwvlieghe AFAD, Rijnders BJA, Philips N, *et al*. Invasive aspergillosis in patients admitted to the intensive care unit with severe influenza: a retrospective cohort study. *Lancet Respir Med* 2018;6:782–92.
19 Nyga R, Maizel J, Nseir S, *et al*. Invasive tracheobronchial aspergillosis in critically ill patients with severe influenza. A clinical trial. *Am J Respir Crit Care Med* 2020;202:708–16.
20 Cornillet A, Camus C, Nimubona S, *et al*. Comparison of epidemiological, clinical, and biological features of invasive aspergillosis in neutropenic and nonneutropenic patients: a 6-year survey. *Clin Infect Dis* 2006;43:577–84.
21 De Pauw B, Walsh TJ, Donnelly JP, *et al*. Revised definitions of invasive fungal disease from the European organization for research and treatment of cancer/invasive fungal infections cooperative group and the National Institute of allergy and infectious diseases Mycoses Study Group (EORTC/MSG) consensus group. *Clin Infect Dis* 2008;46:1813–21.
22 Pfeiffer CD, Fine JP, Safdar N. Diagnosis of invasive aspergillosis using a galactomannan assay: a meta-analysis. *Clin Infect Dis* 2006;42:1417–27.
23 D'Haese J, Theunissen K, Vermeulen E, *et al*. Detection of galactomannan in bronchoalveolar lavage fluid samples of patients at risk for invasive pulmonary aspergillosis: analytical and clinical validity. *J Clin Microbiol* 2012;50:1258–63.
24 Clancy CJ, Jaber RA, Leather HL, *et al*. Bronchoalveolar lavage galactomannan in diagnosis of invasive pulmonary aspergillosis among solid-organ transplant recipients. *J Clin Microbiol* 2007;45:1759–65.
25 Hope WW, Kruhlak MJ, Lyman CA, *et al*. Pathogenesis of Aspergillus fumigatus and the kinetics of galactomannan in an in vitro model of early invasive pulmonary aspergillosis: implications for antifungal therapy. *J Infect Dis* 2007;195:455–66.
26 Horby P, Lim WS, Emberson JR, *et al*. Dexamethasone in hospitalized patients with covid-19. *N Engl J Med* 2021;384:693–704.
27 Koehler P, Cornely OA, Böttiger BW, *et al*. COVID-19 associated pulmonary aspergillosis. *Mycoses* 2020;63:528–34.
28 Alanio A, Dellière S, Fodil S, *et al*. Prevalence of putative invasive pulmonary aspergillosis in critically ill patients with COVID-19. *Lancet Respir Med* 2020;8:e48–9.
29 Salmanton-García J, Sprute R, Stemler J, *et al*. COVID-19-associated pulmonary aspergillosis, march-august 2020. *Emerg Infect Dis* 2021;27:1077–86.
30 Mitaka H, Kuno T, Takagi H, *et al*. Incidence and mortality of COVID-19-associated pulmonary aspergillosis: a systematic review and meta-analysis. *Mycoses* 2021;64:993–1001.
31 Manning J, Dunne EM, Wescombe PA, *et al*. Investigation of streptococcus salivarius-mediated inhibition of pneumococcal adherence to pharyngeal epithelial cells. *BMC Microbiol* 2016;16:225.
32 Jörissen J, van den Broek MFL, De Boeck I, *et al*. Case-control microbiome study of chronic otitis media with effusion in children points at streptococcus salivarius as a pathobiont-inhibiting species. *MSystems* 2021;6:e00056-21.
33 Ilhan ZE, Łaniewski P, Tonachio A, *et al*. Members of Prevotella genus distinctively modulate innate immune and barrier functions in a human three-dimensional endometrial epithelial cell model. *J Infect Dis* 2020;222:2082–92.
34 Picarda G, Benedict CA. Cytomegalovirus: shape-shifting the immune system. *J Immunol* 2018;200:3881–9.