



Clinical Evaluation of a Metagenomics-Based Assay for Pneumonia Management

Yangqing Zhan^{1†}, Teng Xu^{2,3†}, Fusheng He^{2†}, Wei-jie Guan^{1,4†}, Zhengtu Li¹, Shaoqiang Li¹, Mingzhou Xie², Xiaolei Li², Rongchang Chen^{1,5}, Linling Cheng¹, Nanshan Zhong¹ and Feng Ye^{1*}

¹ State Key Laboratory of Respiratory Disease, Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou Institute of Respiratory Health, National Clinical Research Center for Respiratory Disease, Guangzhou, China, ² Vision Medicals Co., Ltd., Guangzhou, China, ³ Key Laboratory of Animal Gene Editing and Animal Cloning in Yunnan Province and College of Veterinary Medicine, Yunnan Agricultural University, Kunming, China, ⁴ Department of Thoracic Surgery, Guangzhou Institute for Respiratory Health, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, ⁵ Department of Respiratory and Critical Care Medicine, First Affiliated Hospital of Southern University of Science and Technology, Second Clinical Medical College of Jinan University, Shenzhen People's Hospital, Shenzhen Institute of Respiratory Diseases, Shenzhen, China

OPEN ACCESS

Edited by:

David Andes,
University of Wisconsin-Madison,
United States

Reviewed by:

Bernard Joseph Hudson,
New South Wales Health Pathology,
Australia
Nicolas Leveque,
University of Poitiers, France

*Correspondence:

Feng Ye
yefeng@gird.cn

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Infectious Diseases,
a section of the journal
Frontiers in Microbiology

Received: 31 July 2021

Accepted: 25 August 2021

Published: 16 September 2021

Citation:

Zhan Y, Xu T, He F, Guan W-J,
Li Z, Li S, Xie M, Li X, Chen R,
Cheng L, Zhong N and Ye F (2021)
Clinical Evaluation of a
Metagenomics-Based Assay
for Pneumonia Management.
Front. Microbiol. 12:751073.
doi: 10.3389/fmicb.2021.751073

Clinical value of metagenomic next-generation sequencing (mNGS) in pneumonia management is still controversial. A prospective study was conducted to evaluate the clinical impact of PneumoSeq in 57 immunocompetent (ICO) and 75 immunocompromised (ICH) pneumonia patients. The value of PneumoSeq for both etiological and clinical impact investigation in pneumonia was assessed. Among the 276 potential pathogens detected with PneumoSeq in our cohort, 251 (90.9%) were cross-validated. Clinical diagnoses of the causative pathogens were obtained for 97 patients, 90.7% of which were supported by PneumoSeq. Compared to conventional testing, PneumoSeq suggested potentially missed diagnoses in 16.7% of cases (22/132), involving 48 additional pathogenic microorganisms. In 58 (43.9%) cases, PneumoSeq data led to antimicrobial treatment de-escalation ($n = 12$ in ICO, $n = 18$ in ICH) and targeted treatment initiation ($n = 7$ in ICO, $n = 21$ in ICH). The PneumoSeq assay benefited the diagnosis and clinical management of both ICH and ICO pneumonia patients in real-world settings.

Keywords: pneumonia, metagenomic next-generation sequence, pathogens, cytomegalovirus, *pneumocystis jirovecii*

INTRODUCTION

Pneumonia is among the most common infectious diseases leading to hospitalization (Prina et al., 2015; Ramirez et al., 2017; Torres et al., 2019) and contributes to significant mortality and morbidity (Barriere, 2010; Torres et al., 2019). According to the World Health Organization (WHO, 2020), pneumonia causes 3 million deaths each year worldwide. The timely and accurate detection of the causative pathogens is crucial for effective tailored clinical management. However, due to the atypical manifestations of pneumonia in immunocompromised hosts (ICHs), empirical management is often ineffective in guiding the appropriate treatment of pneumonia in these patients (Sousa et al., 2013; Azoulay et al., 2019; Di Pasquale et al., 2019). The high prevalence of opportunistic pathogens (e.g., cytomegalovirus (CMV), *Pneumocystis jirovecii*, fungi, *Nocardiosis*)

TABLE 1 | Definition of the study population.

Items	Criteria
Study population	Immunocompromised host (ICH): pneumonia patients who meet the criteria of immunocompromised status Immunocompetent host (ICO): pneumonia patients who do not meet the criteria of immunocompromised status.
Immunocompromised status	Presence of one of the following diseases and/or treatments: <ul style="list-style-type: none"> i. solid-organ or hematopoietic cell transplantation ii. neutropenia iii. solid or hematological malignancy under treatment iv. any type of known immunodeficiency disease v. receiving more than 0.3 mg/kg.d systemic corticosteroids for more than 3 weeks within 60 days or any type of immunosuppressive drugs (such as methotrexate, azathioprine, cyclosporine, or cyclophosphamide)
Inclusion criteria	Community acquired pneumonia, fulfill all clinical features as follow: <ul style="list-style-type: none"> i. community onset ii. a new infiltrate, lobar or segment consolidation, ground glass opacity or interstitial images is revealed on chest radiograph or CT scan, accompanied by pleural effusion or not iii. any of the following 4 clinical features: (I) new occurrence of cough, expectation, or worsen of respiratory symptoms, accompanied by purulent sputum, chest pain, dyspnea and hemoptysis or not; (II) fever; (III) signs of consolidation and/or moist rale on lung auscultation; (IV) peripheral white cell counts $> 10 \times 10^9/L$ or $< 4 \times 10^9/L$. Hospital acquired pneumonia, fulfill all clinical features as follow: <ul style="list-style-type: none"> i. a new or progressive infiltrate, consolidation, or ground glass opacity is revealed on chest radiograph or CT scan ii. two or more of the following 3 criteria: (I) fever $> 38^\circ C$; (II) purulent airway secretions; (III) peripheral white blood cell count $> 10 \times 10^9/L$ or $< 4 \times 10^9/L$ iii. illness occurs 48 hours or more after admission during hospitalization.
Exclusion criteria	<ul style="list-style-type: none"> i. patients infected with HIV ii. pregnant women iii. patients with an irreversible contraindication for bronchoscopy iv. patients who were unable to understand the informed consent description or unwilling to sign the informed consent form v. patients ultimately diagnosed with a non-infectious disease other than pneumonia

and concurrent infections, the difficulty of effectively identifying pathogens based on clinical manifestations, and the limitations of conventional diagnostic techniques have further complicated the effective clinical diagnosis and treatment of pneumonia in ICH patients, leading to failures in the timely administration of targeted antimicrobial drugs. As a result, the mortality rate in ICH patients with pneumonia has remained higher than that in immunocompetent (ICO) patients (Letourneau et al., 2014). Hence, there is an urgent need to develop a diagnostic technique that can rapidly detect a wide range of pathogens in a single test and lead to personalized treatment, especially for immunocompromised host with pneumonia (Chiu and Miller, 2019; Di Pasquale et al., 2019; Azoulay et al., 2020).

Metagenomic next-generation sequencing (mNGS) has emerged as a promising, culture-independent approach for detecting potentially pathogenic microorganisms (Naccache et al., 2014; Simner et al., 2018; Dulanto Chiang and Dekker, 2020; Ramachandran and Wilson, 2020). Unlike assays that target specific genes of pathogens or certain ribosomal RNA amplicons, mNGS is unique in its ability to detect a wide

array of bacteria, viruses, and fungi in an unbiased manner (Naccache et al., 2014; Goldberg et al., 2015; Forbes et al., 2017; Chiu and Miller, 2019; Gu et al., 2019). Data have shown that mNGS can improve the diagnosis of neurologic infections and provide actionable information in some cases (Wilson et al., 2019). However, the application of mNGS for respiratory pathogen detection is challenging due to the wide variety and high abundance of microbes within airways. Fungal infection is also common among ICH patients (Di Pasquale et al., 2019). In theory, mNGS, which take a short turnout time, and detect all microorganisms at one time, is an ideal detection method for patients with ICH pneumonia to determine the etiology as soon as possible and to carry out targeted individualized treatment. Nevertheless, there is currently a lack of prospective studies assessing the use of mNGS for guiding the etiology diagnosis and treatment of pneumonia in ICH patients (Zinter et al., 2019a).

We hypothesized that an optimized mNGS assay could improve the clinical management of pneumonia. Our study aimed to better understand the microbial etiology of pneumonia in ICO and ICH patients and the value of

TABLE 2 | Evaluation guidance for role of microorganisms detected by PneumoSeq.

Classification of microorganisms	Role of microorganisms
Common pathogens of pneumonia: <i>Streptococcus pneumoniae</i> , <i>Haemophilus influenzae</i> , <i>Moraxella catarrhalis</i> , <i>Klebsiella pneumoniae</i> , <i>Escherichia coli</i> , <i>Acinetobacter baumannii</i> , <i>Pseudomonas aeruginosa</i> , <i>Staphylococcus aureus</i> , et al.	Attention was paid to these microorganisms if they were positive by PneumoSeq. They were considered components of the microbiological etiology of pneumonia if they were also positive in culture.
Atypical pathogens: <i>Mycoplasma pneumoniae</i> , <i>Chlamydia pneumoniae</i> , <i>Legionella pneumophila</i>	They were considered components of the etiology of pneumonia if the index patient met the clinical criteria of infection, whether confirmed by conventional tests or not.
Special bacteria: <i>Nocardia</i> , <i>Mycobacterium tuberculosis</i> , <i>Non-tuberculosis mycobacteria</i>	They were considered components of the etiology of pneumonia if the index patient met the clinical criteria of infection, whether confirmed by conventional tests or not.
Respiratory microecological bacteria	They were mostly considered normal flora
Fungus (<i>Candida</i> is excluded because primary pulmonary <i>Candida</i> pneumonia is rare)	Attention was paid to fungi. They were considered components of the etiology of pneumonia if the index patient fulfilled risk factor and clinical criteria of pulmonary fungal disease by EORTC/MSG and responded to anti-fungal treatment, whether accompanied by microbiological or cytological evidence or not.
Respiratory virus: Influenza A/B virus, parainfluenza-1/2/3, adenovirus, respiratory syncytial virus, human bocavirus, human metapneumovirus, human coronavirus 229E/OC43/HKU, human rhinovirus, Cytomegalovirus, EB virus	They were considered components of the etiology of pneumonia if the index patient met the clinical criteria of infection, whether confirmed by conventional tests or not.
Microorganisms with unclear significance: Torque teno virus, Torque teno mini virus, KI polyomavirus, et al.	They were considered as non-pathogenic.

TABLE 3 | Guidelines for evaluating the clinical benefits of PneumoSeq.

Clinical benefit of PneumoSeq	Guidance from PneumoSeq in the management of patients
Initiation of targeted treatment	PneumoSeq detected pathogens and guided targeted antimicrobial treatment.
Pathogen identification or treatment confirmation	PneumoSeq detected pathogens, either consistent with conventional tests or not, and/or PneumoSeq confirmed that the initial empirical treatment could be continued.
Treatment De-escalation	Antimicrobial drugs were de-escalated or discontinued according to the results of PneumoSeq.
No clinical benefit	Antimicrobial drugs were adjusted according to conventional tests or physician experience.

PneumoSeq in the clinical management of pneumonia (especially in ICH patients).

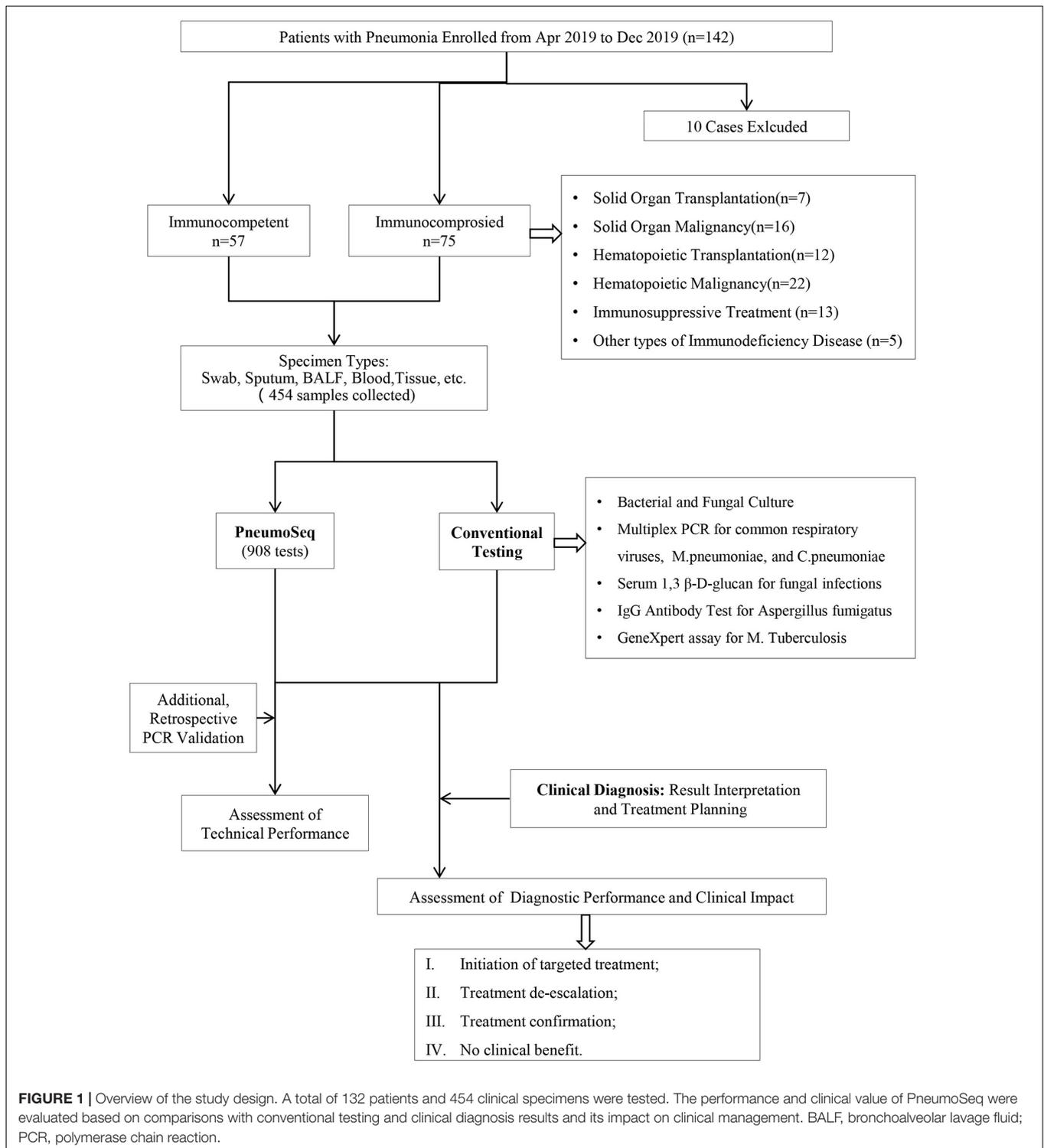
MATERIALS AND METHODS

Patients and Samples

This prospective study was conducted in the Department of Pulmonary and Critical Care Medicine, the First Affiliated Hospital of Guangzhou Medical University, between April 2019 and January 2020. Consecutive patients older than 14 years with community acquired pneumonia (Cao et al., 2018) or hospital acquired pneumonia (Shi et al., 2019) were enrolled and divided into ICH or ICO group according to the immune status of the patients, with details shown in **Table 1**. This study was approved by the ethics committee of the First Affiliated

Hospital of Guangzhou Medical University (No. 2019-49), and written informed consent was obtained from all subjects or their guardians.

All patients received conventional diagnostic methods for identification of pathogens based on the patient's condition during hospitalization. The choice of conventional diagnostic methods was determined by physician and conventional diagnostic methods could be repeated if necessary. The most common conventional diagnostic methods for microbiological diagnosis were culture for bacteria and fungi, polymerase chain reaction (PCR)-based assays for common respiratory viruses and mycoplasma pneumoniae, and serum tests for fungal infections. Specimen types included throat swabs, sputum, bronchoalveolar lavage fluid (BALF), blood, and pleural effusion, etc. Throat swab was only used for detection of respiratory viruses and atypical pathogens via PCR, including influenza virus type A and B,



parainfluenza virus type 1/2/3, adenovirus, respiratory syncytial virus, mycoplasma pneumoniae, chlamydia pneumoniae and legionella pneumophila. Surplus of the sputum and BALF after conventional diagnostic methods was tested by PneumoSeq. In addition, throat swabs and blood were also collected and

tested by PneumoSeq to match sample types of the conventional diagnostic methods. Another purpose of selection of throat swabs for PneumoSeq was to compare the role of different specimen types in pathogenic diagnosis. Additional PCR assays were retrospectively conducted for the validation of the positive

TABLE 4 | Demographic data and underlying conditions of immunocompromised and immunocompetent patients with pneumonia.

Clinical features	ICH	ICO	P value
Number of cases	75	57	NA
Sex, female	28 (37%)	28 (49%)	0.214
Age (mean ± sd)	50 ± 19	43 ± 20	0.068
Solid organ transplantation	7 (9%)	NA	NA
Hematopoietic cell transplantation	12 (16%)	NA	NA
Solid-organ malignancy	16 (21%)	NA	NA
Hematopoietic malignancy	22 (29%)	NA	NA
Immunosuppressive treatment	13 (17%)	NA	NA
Other types of immunodeficiency disease	5 (7%)	NA	NA
CTD	3 (4%)	0 (0%)	0.258
COPD	3 (4%)	5 (9%)	0.29
Hypertension	15 (20%)	7 (12%)	0.346
Diabetes	5 (7%)	4 (7%)	1
Chronic hepatic disease	8 (11%)	2 (4%)	0.186
Chronic renal disease	10 (13%)	3 (5%)	0.149
Antibiotic 30 days before onset of pneumonia	15 (20%)	0 (0%)	<0.001
Smoker	23 (31%)	14 (25%)	0.558
Liquor	8 (11%)	3 (5%)	0.349
White blood cell counts (mean ± sd)	8.5 ± 5	9.6 ± 6.3	0.27
Neutrophil [%] (mean ± sd)	73 ± 19	86 ± 95	0.308
Lymphocyte [%] (mean ± sd)	17 ± 16	16 ± 9.5	0.781
ESR (mean ± sd)	66 ± 36	79 ± 33	0.092
C-reactive protein (mean ± sd)	10 ± 22	15 ± 28	0.471
Procalcitonin (mean ± sd)	3.2 ± 21	0.8 ± 2.5	0.342
ICU admission	6 (8%)	2 (4%)	0.465
Death within 30 days	3 (4%)	1 (2%)	0.816

NA, the clinical features were not suitable for the control group; CTD, connective tissue diseases; COPD, chronic obstructive pulmonary disease; ESR, erythrocyte sedimentation rate; ICU, intensive care unit.

PneumoSeq results if the index patient didn't receive PCR testing during hospitalization.

Sample Preparation, Library Construction and Sequencing for PneumoSeq

Whole-blood samples were collected in Geneseek serum cell-free DNA tubes and transported at ambient temperature. Other specimens were stored and transported at 4–8°C after collection. Prior to extraction, BALF and homogenized sputum specimens were centrifuged at 8,000 *g* for 5 min before being resuspended in lysis buffer with 0.1% sodium dodecyl sulfate (SDS) and 1% Nonidet P-40 (NP40). The dried oropharyngeal swab samples were soaked in 1 ml of sterile PBS for 10 min. Plasma was separated from whole blood by centrifugation at 500 *g* for 5 min. For DNA extraction, 600 µL of the processed specimens was mixed with glass beads of 0.1–0.2 mm diameter. A vortex mixer (Crystal, TX, United States) was used to disrupt the bacterial cell wall at 1,600 *g* for 10 min. The tubes were then heated at 99°C for 10 min before DNA extraction. RNA extraction was performed in parallel. The concentration of extracted DNA/RNA was measured using a Qubit Fluorometer before library preparation.

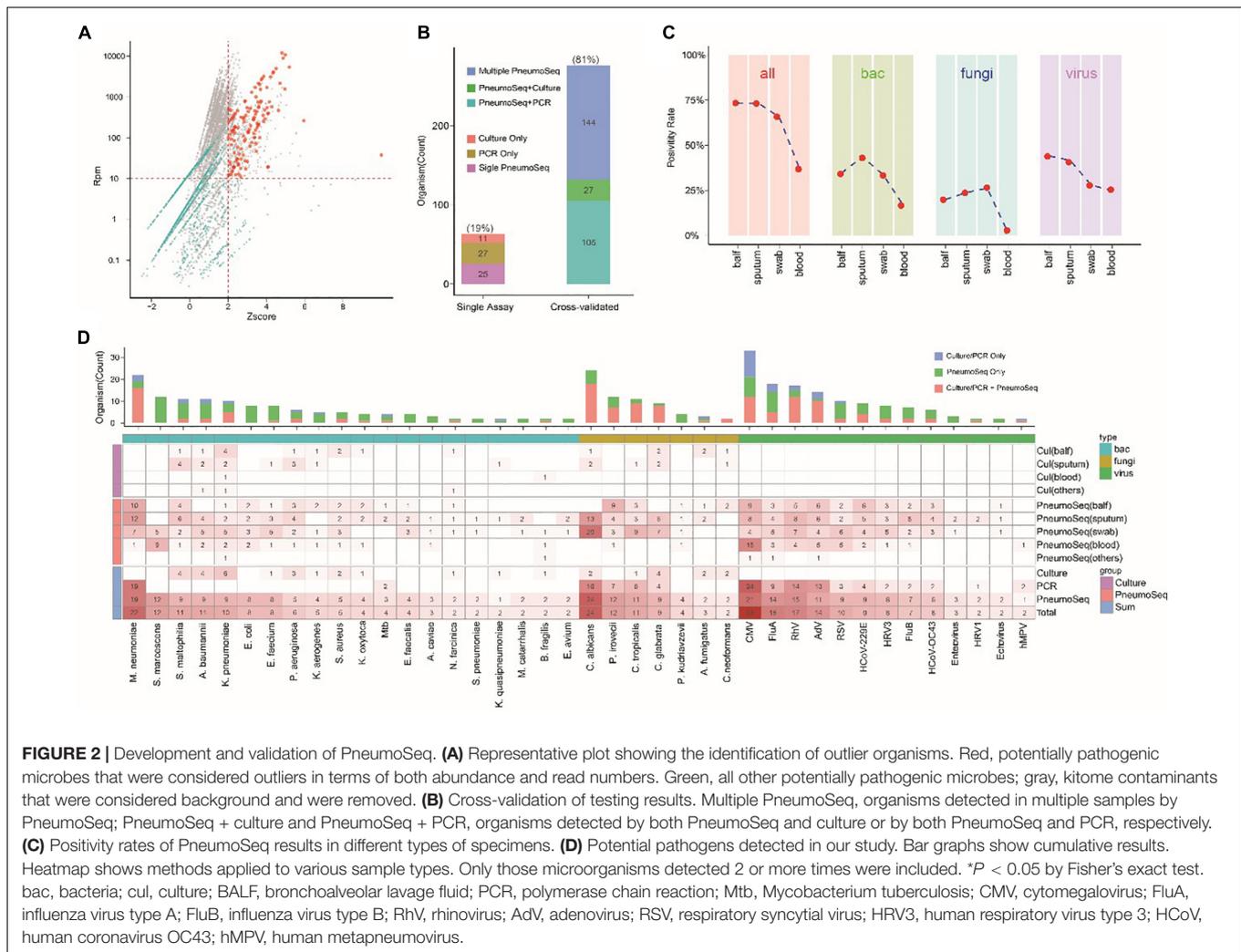
DNA libraries were prepared via transposase-based methodology. Human rRNA was depleted from the RNA samples

via an RNase H-based method before library preparation. After purification and size selection, the concentration of the RNA library was determined by using a Qubit instrument before pooling. Pooled libraries were sequenced on an Illumina NextSeq 550 system using a 75 bp, single-end sequencing kit (Illumina, San Diego). The qualified results had no fewer than 15 million reads obtained per sample and a Q30 score of 90% or greater. A negative control sample was processed and sequenced in parallel in each sequencing run for quality control.

Bioinformatic Pipeline for PneumoSeq

The 75bp single-end reads from illumine Nextseq 550 were analyzed by in-house IDseq software to get each microorganism's abundance. The detail process is as follows: high-quality sequencing data were generated by removing reads of low quality or short length (<35 bp) by using fastp (Chen et al., 2018). Human host sequences were subtracted by mapping to human reference genome sequences (National Center for Biotechnology Information GRCh38 assembly) using the Burrows-Wheeler Aligner tool (BWA)¹ (Li and Durbin, 2009). The data remaining after the removal of low-complexity reads were classified by alignment to curated microbial genome databases for

¹<http://bio-bwa.sourceforge.net/>



viruses, bacteria, fungi, and parasites. Taxonomic references were downloaded from the National Center for Biotechnology Information.²

After each microorganism is quantified, it is necessary to remove the contamination from the reagent (kitome). To build the background filter for kitome removal, we categorized only the microorganisms that were detected in at least 25% of the samples, including the negative controls. Assuming that the kitome contaminants existed at a fixed abundance, the percentage of contamination would depend on and negatively correlate with the amount of extracted DNA/RNA. Pearson's and Spearman's correlations were used for the correlation analyses. Species that showed significant negative correlations between the abundance inferred from normalized read numbers and the extracted concentration of nucleic acids were considered reagent-derived background organisms and therefore excluded from reporting.

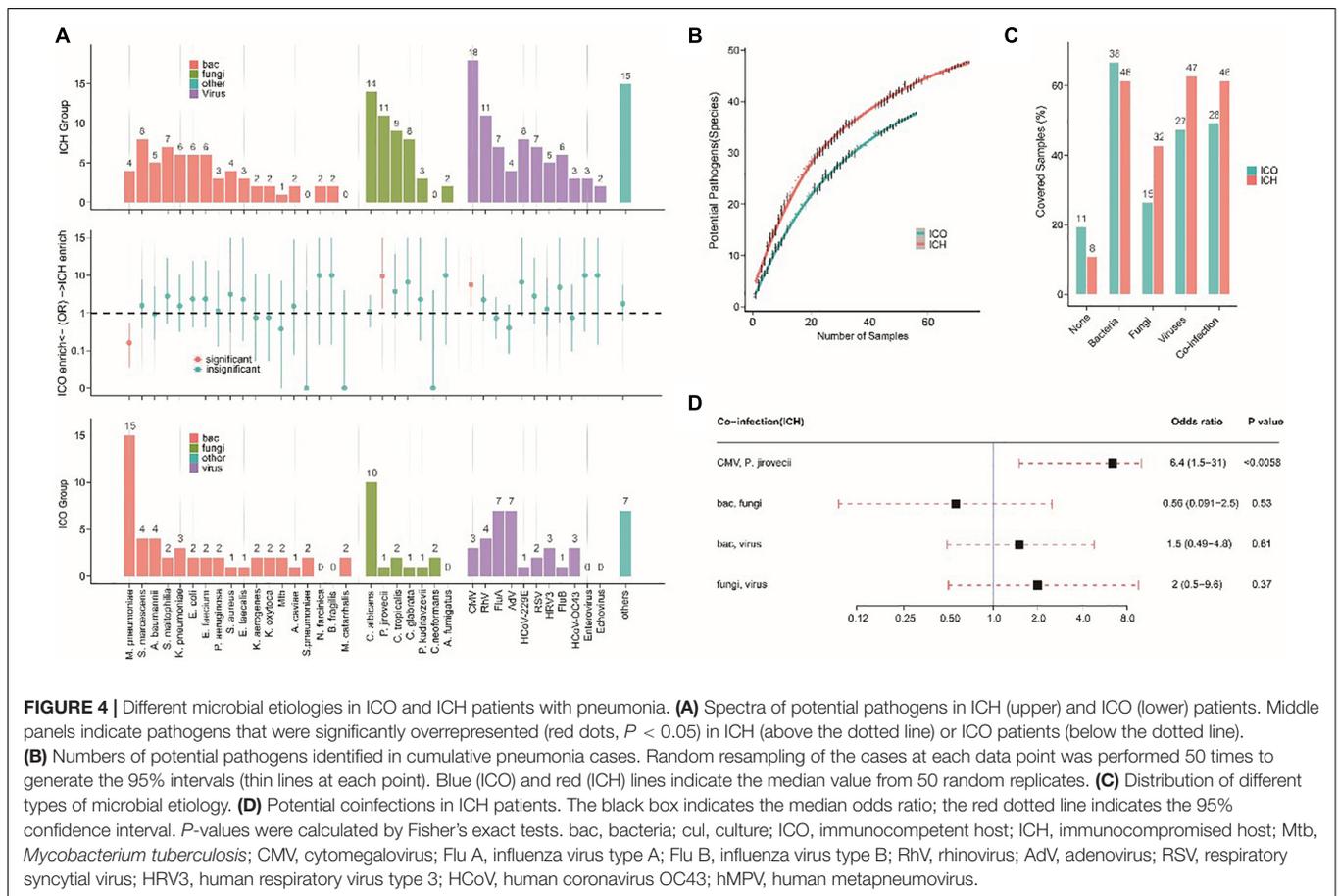
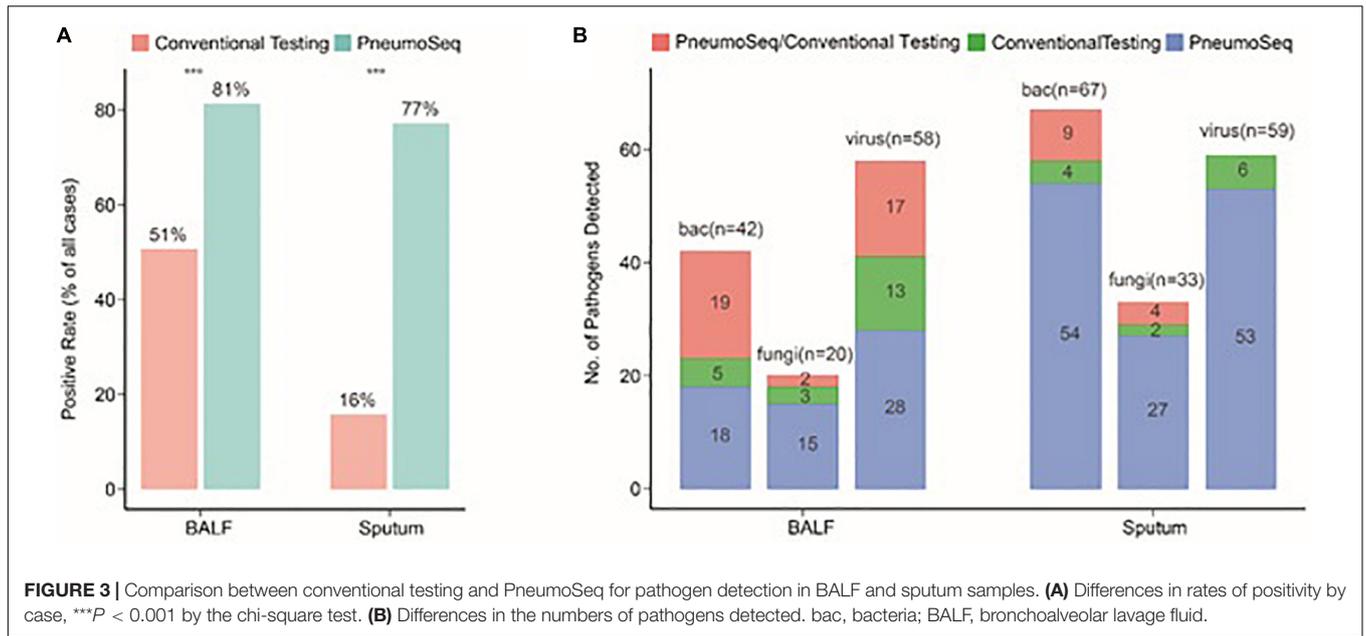
The conditions for selection and reporting vary for different types of microorganisms. In the case of bacteria and fungi, only

outliers need to be considered, and in the case of viruses, reads need to be cover multiple areas. Microbial outliers were defined on the basis of both read-number-based and Z-score-based abundance. To minimize read-depth variation, the actual read numbers of each identified organism were normalized to reads per million total reads (RPM). The Z-score statistic was employed to identify the outliers of each organism for which abundance significantly deviated from that of the total cohort. We defined microbial outliers as those with Z-scores ≥ 2 and ≥ 10 RPM (bacteria), or ≥ 3 reads from distinct genomic regions (viruses), or ≥ 2 RPM (fungi) (Miller et al., 2019; Zinter et al., 2019a).

Clinical Assessment

The results of PneumoSeq were reviewed along with other clinical evidence by a group of senior physicians. The clinical significance of the microorganisms detected by PneumoSeq was analyzed based on the clinical characteristics of the patient and the lung pathogenicity of the identified microorganisms (Table 2). In summary, attention was paid for common bacteria pathogens of pneumonia, such as *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Moraxella catarrhalis*, *Klebsiella pneumoniae*,

²ftp://ftp.ncbi.nlm.nih.gov/genomes/



Escherichia coli, *Acinetobacter baumannii*, *pseudomonas aeruginosa*, *Staphylococcus aureus*, et al., if they were positive by PneumoSeq. They were considered components of the

microbiological etiology of pneumonia if they were also positive in culture. When atypical pathogens, respiratory virus or pathogenic bacteria was positive by PneumoSeq, they were

TABLE 5 | Microbial spectra differ between ICH and ICO patients with pneumonia.

Species	Type	ICO(57)	ICH(75)	Fisher's exact p value	Odds ratio (95% CI)
<i>Mycoplasma_pneumoniae</i>	Bac	15	4	< 0.001	0.16 (0.04–0.55)
<i>Serratia_marcescens</i>	Bac	4	8	0.552	1.58 (0.40–7.55)
<i>Acinetobacter_baumannii</i>	Bac	4	5	1	0.95 (0.19–5.01)
<i>Stenotrophomonas_maltophilia</i>	Bac	2	7	0.298	2.81 (0.51–28.80)
<i>Klebsiella_pneumoniae</i>	Bac	3	6	0.731	1.56 (0.32–10.08)
<i>Escherichia_coli</i>	Bac	2	6	0.465	2.38 (0.40–24.98)
<i>Enterococcus_faecium</i>	Bac	2	6	0.465	2.38 (0.40–24.98)
<i>Pseudomonas_aeruginosa</i>	Bac	2	3	1	1.14 (0.13–14.14)
<i>Staphylococcus_aureus</i>	Bac	1	4	0.389	3.13 (0.30–157.98)
<i>Enterococcus_faecalis</i>	Bac	1	3	0.633	2.32 (0.18–124.57)
<i>Klebsiella_aerogenes</i>	Bac	2	2	1	0.76 (0.05–10.71)
<i>Klebsiella_oxytoca</i>	Bac	2	2	1	0.76 (0.05–10.71)
<i>Mycobacterium_tuberculosis</i>	Bac	2	1	0.578	0.37 (0.01–7.36)
<i>Aeromonas_caviae</i>	Bac	1	2	1	1.53 (0.08–92.06)
<i>Streptococcus_pneumoniae</i>	Bac	2	0	0.185	0.00 (0.00–4.03)
<i>Nocardia_farcinica</i>	Bac	0	2	0.506	Inf
<i>Bacteroides_fragilis</i>	Bac	0	2	0.506	Inf
<i>Moraxella_catarrhalis</i>	Bac	2	0	0.185	0.00 (0.00–4.03)
<i>Candida_albicans</i>	Fungi	10	14	1	1.08 (0.40–2.97)
<i>Pneumocystis_jirovecii</i>	Fungi	1	11	0.013	9.51 (1.31–420.71)
<i>Candida_tropicalis</i>	Fungi	2	9	0.113	3.72 (0.73–36.75)
<i>Candida_glabrata</i>	Fungi	1	8	0.077	6.61 (0.84–301.47)
<i>Pichia_kudriavzevii</i>	Fungi	1	3	0.633	2.32 (0.18–124.57)
<i>Cryptococcus_neoformans</i>	Fungi	2	0	0.185	0.00 (0.00–4.03)
<i>Aspergillus_fumigatus</i>	Fungi	0	2	0.506	Inf
CMV	Virus	3	18	0.004	5.62 (1.52–31.46)
Rhinovirus	Virus	4	11	0.268	2.26 (0.62–10.32)
Influenza_A_virus	Virus	7	7	0.584	0.74 (0.21–2.64)
Human_adenovirus	Virus	7	4	0.206	0.41 (0.08–1.69)
Human_coronavirus_229E	Virus	1	8	0.077	6.61 (0.84–301.47)
Human_respiratory_syncytial_virus	Virus	2	7	0.298	2.81 (0.51–28.80)
Human_respirovirus_3	Virus	3	5	1	1.28 (0.24–8.62)
Influenza_B_virus	Virus	1	6	0.140	4.82 (0.56–227.67)
Human_coronavirus_OC43	Virus	3	3	1	0.75 (0.10–5.83)
Enterovirus	Virus	0	3	0.258	Inf
Echovirus	Virus	0	2	0.506	Inf
Other	Other	7	15	0.346	1.78 (0.62–5.58)

ICH, immunocompromised host; ICO, immunocompetent host; CMV, cytomegalovirus.

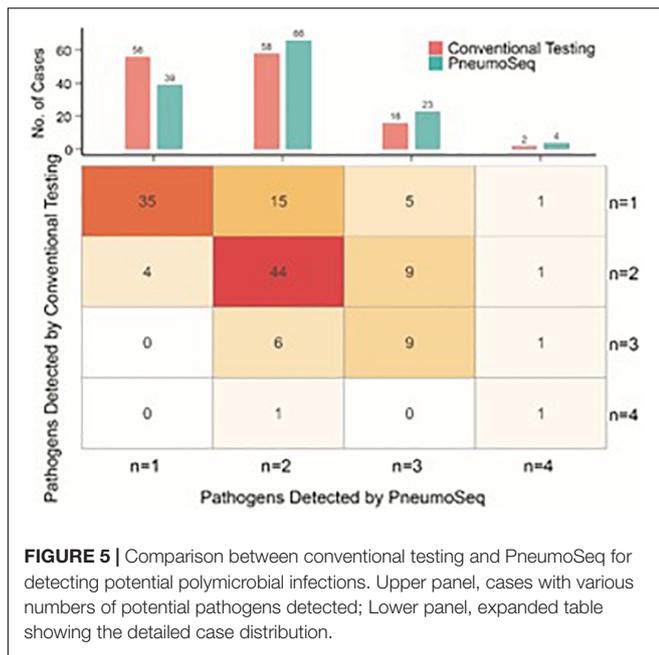
considered components of the etiology of pneumonia if the index patient met the clinical criteria of infection, whether confirmed by conventional tests or not. Attention was paid to most fungi other than yeast. They were considered components of the etiology of pneumonia if the index patient fulfilled risk factor and clinical criteria of pulmonary fungal disease by EORTC/MSG and responded to anti-fungal treatment, whether accompanied by microbiological or cytological evidence or not.

Changes in treatment plans were made after a positive result of PneumoSeq if indicated. It should be noted that the results of throat swabs by PneumoSeq were only considered for diagnosis of respiratory viruses and/or atypical pathogens infections, but not suitable for the bacteria and fungi. After discharge, the microbiological etiology of each case and the clinical impact of PneumoSeq were assessed by

the same senior physicians according to the clinical data and the guidance for the clinical benefit evaluation of PneumoSeq, which was classified as initiation of targeted treatment, pathogen identification or treatment confirmation, treatment de-escalation or no clinical benefit (**Figure 1** and **Table 3**).

Statistical Analysis

Comparative analyses were conducted by using the chi-square test, Fisher's exact test for continuous variables, and Student's *t*-test or Wilcoxon rank sum test for numerical variables where appropriate. Data analyses were performed using R software. *P* values <0.05 were considered significant, and all tests were 2-tailed unless otherwise indicated.



RESULTS

Patient Characteristics and Samples

We enrolled 142 patients with pneumonia, 132 of whom completed the study (Figure 1). The median age of the 132 patients was 45 years, ranging from 14 to 83. Fifty-seven patients were immunocompetent, and seventy-five patients were immunocompromised (Figure 1 and Table 4); hematological malignancy (29%) was the most common comorbidity, followed by solid-organ malignancy (21%), immunosuppressive treatment (17%) and hematopoietic cell transplantation (16%). Twenty-eight patients were severely ill as defined by the APACHE II system, among which 6 ICHs (8%) and 2 ICOs (4%) were admitted to the ICU during hospitalization. The 30-day mortality among all study patients was 3% (Table 4). A total of 435 specimens were collected. Overall, specimens were collected from each patient, among which bronchoalveolar lavage fluid (BALF, $n = 75$), sputum ($n = 101$), plasma ($n = 124$), and nasopharyngeal swabs ($n = 122$) were the most common types. All patients were tested with PneumoSeq and conventional testing.

Development and Validation of PneumoSeq

The performance of the mNGS assay for diagnosing pneumonia involving various microorganisms (especially fungi with thicker cell walls) is an important metric reflecting the clinical value. Hence, we improved the assay sensitivity by establishing an effective extraction method based on a combination of bead beating, chemical lysis, and heating to ensure efficient nucleic acid extraction. As experimental reagents can cause false-positive results by introducing kitome microbes (Zinter et al., 2019b), we developed an algorithm for filtering these background species and integrated it into the PneumoSeq assay

to reduce false-positive results. Furthermore, a Z-score model was employed to identify outlier microbes to achieve high assay specificity (Figure 2A).

PneumoSeq was performed on a total of 435 clinical samples. Among the 339 potential pathogens identified by PneumoSeq and/or conventional testing, 81% were cross-validated by using either another assay or other samples collected from the same patient (Figure 2B). When considering only the potential pathogens detected with PneumoSeq, we observed a cross-validation rate of 90.9% (251/276), confirming the reliability of the optimized assay.

To identify the most suitable specimen type for PneumoSeq in patients with pneumonia, we compared the positivity among four major sample types. BALF and sputum were the most effective sample types for pathogen detection, showing a positivity rate of 75%, whereas only 35% of the plasma samples yielded positive results (Figure 2C). The results for nasopharyngeal swabs appeared to be comparable to those for sputum and BALF, except that a lower rate of viral detection was observed. Relative to conventional testing, the application of PneumoSeq led to significantly higher rates of pathogen detection in BALF (81 vs 51%, $P < 0.005$) and sputum samples (77 vs 16%, $P < 2.894e-12$), which was mostly attributed to the improved detection of bacterial and viral pathogens (Figure 3). As a single diagnostic assay, PneumoSeq covered a broad microbial spectrum and detected infections that were readily missed by conventional detection methods (Figure 2D).

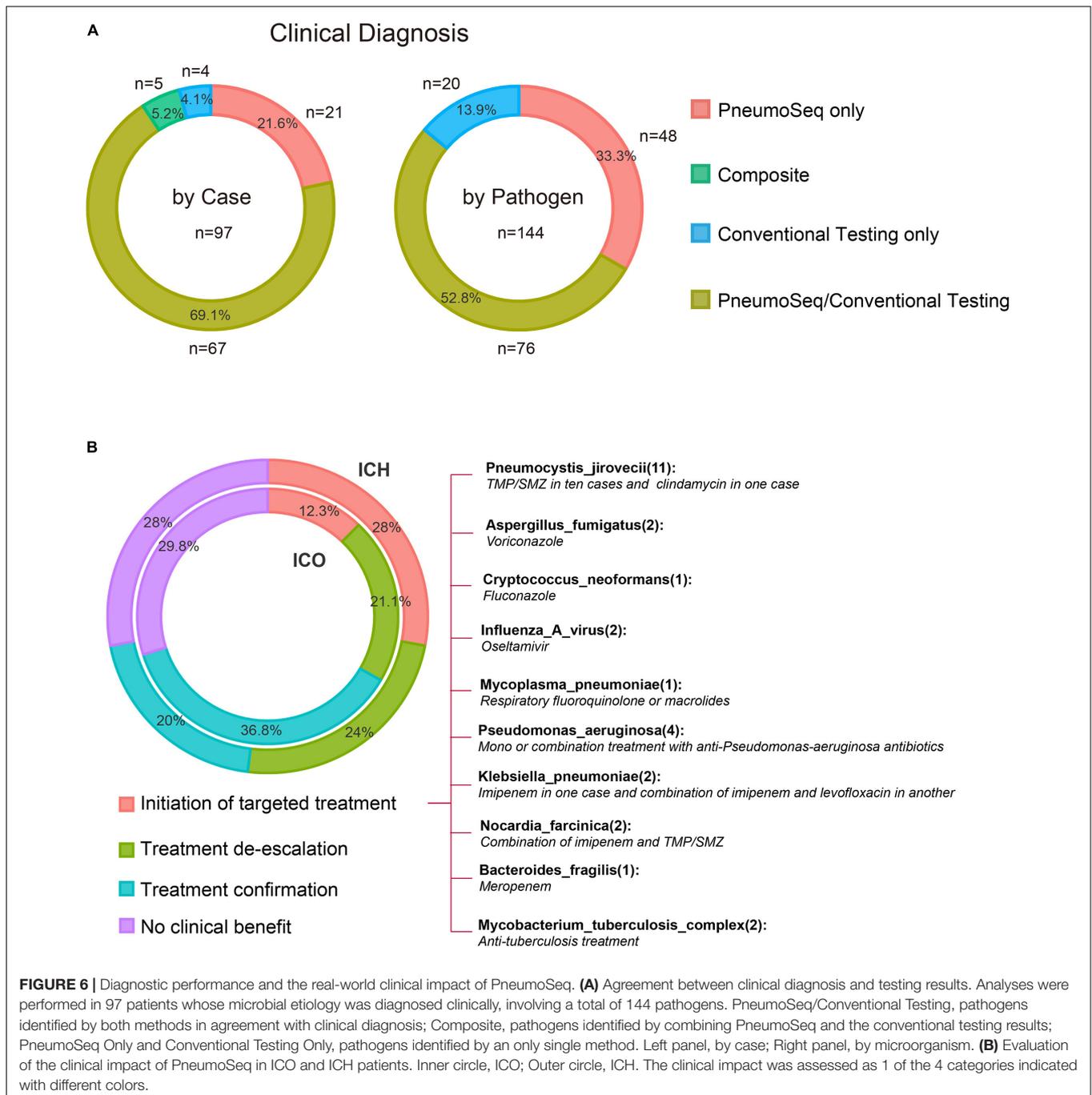
Microbial Spectra Differ Between ICH and ICO Patients With Pneumonia

To further evaluate the microbial etiology of pneumonia, we compared the potential pathogens between the ICH and ICO groups. We found a differential microbial spectrum in ICH patients with pneumonia (Figure 4A). These patients exhibited significantly higher rates of *Pneumocystis jirovecii* (OR = 9.5, 95% CI, 1.3–420.7, $P = 0.01$) and CMV (OR = 5.6, 95% CI, 1.5–31.5, $P < 0.01$) infections than ICO patients, while the rate of *Mycoplasma pneumoniae* infections was significantly lower in ICH patients (OR = 0.16, 95% CI, 0.04–0.55, $P < 0.001$, Table 5).

The simultaneous detection of two or more pathogens was more common in ICH patients with pneumonia, with a significantly greater number of potential pathogens being identified in each patient than in their ICO counterparts (2.7 vs 1.8, $P < 0.01$, Figure 5). The overall diversity of pathogens was also greater in ICH patients with pneumonia (Figure 4B, $P < 0.01$). When grouped by microbial type, ICH patients with pneumonia showed a higher percentage of viral, fungal, and polymicrobial infections but a slightly lower percentage of bacterial infections (Figure 4C). Moreover, *P. jirovecii* infections were more significantly overrepresented in ICH patients with CMV (OR = 6.4, 95% CI, 1.5–31.0, $P < 0.01$, Figure 4D).

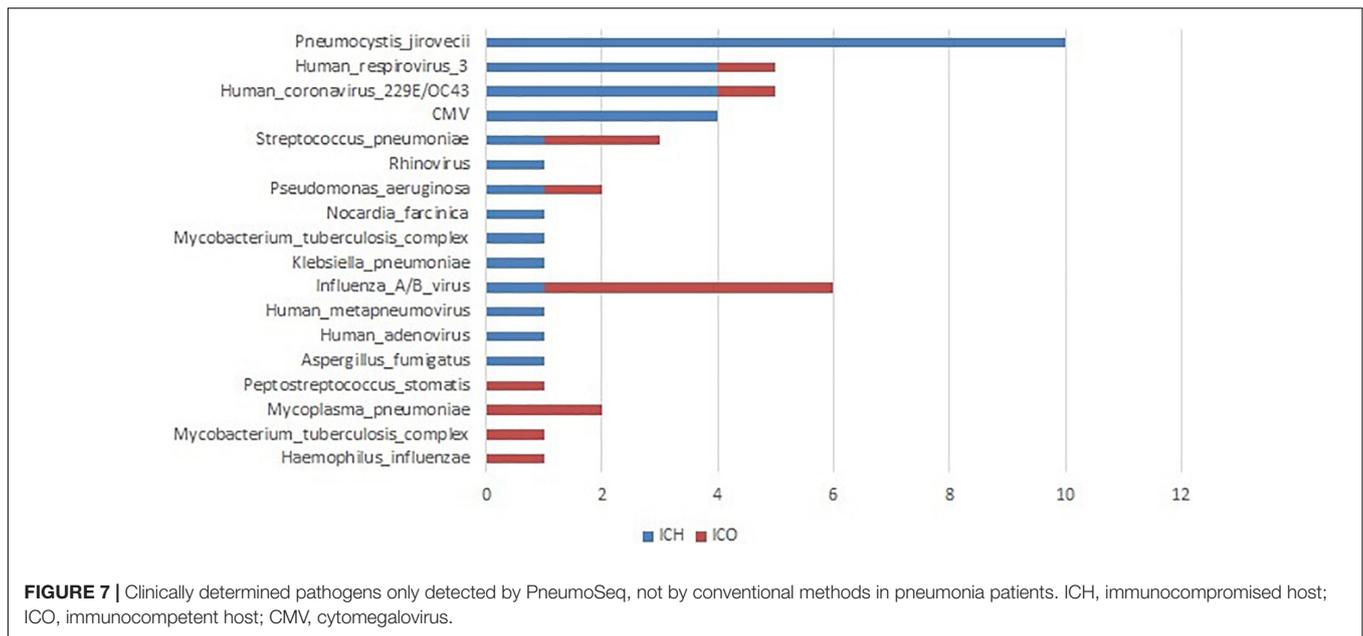
Assessment of the Real-World Clinical Impact of PneumoSeq

All patients underwent both conventional and PneumoSeq testing. Among the 97 patients who received clinical diagnoses



of the causative pathogens (73.5%, 97/132), 90.7% (88/97) of the diagnoses were supported by PneumoSeq. Conventional testing yielded a missed diagnosis in 16.7% of the cases (22/132), whereas conventional testing detected pathogens in 2.3% (3/132) of the cases that were deemed negative by PneumoSeq. Similarly, among all clinically diagnosed pathogens, 86.1% (124/144) showed findings consistent with those of PneumoSeq. Importantly, PneumoSeq reported an additional 48 pathogenic microorganisms that could have been missed if using only conventional methods (Figure 6A and Figure 7).

We finally attempted to assess the clinical impact of the assay's findings. PneumoSeq led to an overall beneficial impact on treatment in 33.3% (19/57) of the ICO patients and 52.0% (39/75) of the ICH patients. The beneficial impacts were categorized as either the successful de-escalation of treatment ($n = 12$ in ICO, $n = 18$ in ICH) or the initiation of targeted treatment ($n = 7$ in ICO, $n = 21$ in ICH). PneumoSeq also helped identify the etiology or confirm appropriate treatment in 36.8% (21/57) and 20.0% (15/75) of the ICO and ICH patients, respectively. PneumoSeq showed no clinical benefit because of failure in identifying



pathogens or guiding targeted antimicrobial treatment in slightly less than 30% of both ICO and ICH patients (Figure 6).

DISCUSSION

By enrolling both ICO and ICH patients with pneumonia and testing for multiple types of specimens in the same individual patient, we evaluated the value of PneumoSeq for the diagnosis of pneumonia and assessed its clinical impact in a real-world setting in both ICO and ICH patients. We found that PneumoSeq detected additional pathogens in 22 patients, including CMV, *P. jirovecii*, *Nocardia*, and fungi in the ICH group and influenza virus, *M. pneumoniae*, and *S. pneumoniae* in the ICO group, lead to de-escalation of antimicrobial treatment and initiation of appropriate treatment.

The rate of confirmed etiologies in our study was higher than that previously reported (77 vs 44% in ICH, 77 vs 32% in ICO) (Sousa et al., 2013), with a considerable proportion of polymicrobial infections (40% in ICH and 15.8% in ICO). PneumoSeq identified potential pathogens in 70.5% of the cases. Missed infections were suggested in 22 patients with negative results according to all other conventional assays. The additional pathogens detected by PneumoSeq included CMV, *P. jirovecii*, *Nocardia*, and fungi in the ICH group and influenza virus, *M. pneumoniae*, and *S. pneumoniae* in the ICO group. The reasons for false negative of these pathogens by traditional diagnostic methods included lack of detection technology, such as authenticated pneumocystis nucleic acid detection assay, low positivity in pathology or cytology for fungi, *P. jirovecii*, etc., low positivity in culture for *Nocardia*, *P. streptococcus*, etc. Moreover, some pathogens such as CMV, *M. pneumoniae*, etc. was not considered by clinicians during hospital admission based

on empirical judgments, and therefore conventional assays were not performed.

For the above-mentioned pathogens additionally detected by PneumoSeq, targeted drugs are available for clinician to make personalized treatment for patients. Hence, the improved ability to identify the microbial etiology by using PneumoSeq further enabled the application of targeted treatment in 12.3% of ICO patients and 28% of ICH patients. These results suggest the value of PneumoSeq in guiding the clinical management of pneumonia, especially in ICH patients. The clinical benefit of mNGS observed in our study was superior to its reported utility in a prospective study for diagnosing meningitis (14.3%) (Wilson et al., 2019) and that of plasma cell-free DNA in a retrospective study for exploring the pathogens of infectious diseases (7.3%) (Hogan et al., 2020). There was a lack of prospective studies in the literature on the application of mNGS in lung infectious diseases. The existing data were basically retrospective studies, exploring the value in pathogenic diagnosis. There were limited data to assess whether the patients will benefit from the result of mNGS (Chen et al., 2021; Peng et al., 2021; Zhang et al., 2021).

Our data also showed a different spectrum of pathogens in ICH patients with pneumonia than in their ICO counterparts. ICH patients had a more complex microbial etiology, with higher detection rates of fungal, viral and coexisting infectious agents, such as CMV and *P. jirovecii*. These findings emphasized the challenges in the diagnosis of ICHs with pneumonia by conventional testing. PneumoSeq improved yield of pathogens and lead to personalized targeted antimicrobial treatment in ICH patients.

Kalantar et al. (2019) applied mNGS to compare the yield of tracheal aspirate and minibronchoalveolar lavage (mBAL) specimens in the assessment of the respiratory microbiota of ventilated patients, and they found that the use of tracheal aspirates provided a similar assessment of

airway microbiota to that obtained using mBAL in patients with pneumonia (Kalantar et al., 2019). By comparing the PneumoSeq results from tests of different specimen types, BALF and sputum samples were found to be the most appropriate sample types for pathogen detection, with BALF showing less contamination by oral microorganisms. In patients whose BALF is unavailable at a particular time, sputum can serve as a promising alternative. Although nasopharyngeal swabs show a slightly lower positivity rate, they could serve as a practical alternative sample type when neither of the other two types is available.

Although PneumoSeq showed a higher diagnostic performance than other conventional methods, as a high-throughput sequencing-based assay, its sensitivity could be affected by the abundance of pathogens and the host cell content in each sample. While PneumoSeq identified more pathogens than other conventional methods combined, potential false-negative results were also observed, as suggested by other assays. This could be due to the high abundance of host cells in these samples (Chiu and Miller, 2019; Miller et al., 2019), which limited the detection ability of PneumoSeq to a greater degree than that of other targeted methods, such as PCR. This should be noted when interpreting results from any metagenomic-based assay.

As a promising diagnostic tool, the wider application of mNGS is still hindered by its cost and turnaround time (Flygare et al., 2016). Currently, an mNGS assay costs \$1000–2000 in the US and \$400–500 in China (Miao et al., 2018; Ramachandran and Wilson, 2020), and the turnaround time is usually 24–48 h (Simner et al., 2018). However, it is cheaper than the total conventional microbiological testing cost of approximately \$1000. As the network of clinical laboratories that offer this new assay expands, its affordability and feasibility are expected to be improved in clinical settings (Simner et al., 2018; Han et al., 2019).

mNGS is unique in that it collects unbiased sequence information of pathogens as well as host genes (Avraham et al., 2016; Chiu and Miller, 2019). Some previous studies have shown that this information could be used for analyses of drug resistance and virulence genes and even host responses and that it offers added diagnostic value (Fournier et al., 2014; Forbes et al., 2017; Chiu and Miller, 2019). More studies are still needed for the validation of these assays and the assessment of their clinical impact.

Our study has limitations as a single-center study in which all patients were enrolled from a single institution. Some patients had received ineffective treatment prior to being referred to our hospital. This might have resulted in a decreased positivity

rate by culture and a higher chance of infections caused by opportunistic pathogens.

In conclusion, our study has demonstrated the clinical value of PneumoSeq for pathogen detection and management in ICH and ICO patients with pneumonia.

DATA AVAILABILITY STATEMENT

The data can be found in the China National GeneBank DataBase (CNGBdb), which is a unified platform for the scientific research community to provide biological big data sharing and application services (<https://db.cngb.org/search/project/CNP0002153/>).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the First Affiliated Hospital of Guangzhou Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

YZ, TX, FH, and MX contributed to conception and design of the study. ZL and SL organized the database. FH, XL, and MX performed the statistical analysis. YZ and FH wrote the first draft of the manuscript. TX and W-JG wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by National Major Science and Technology Projects: 2017ZX09304005002 and Guangdong Province Major Area Research and Development Projects: 2018B020241002.

ACKNOWLEDGMENTS

We acknowledge Ye Lin for her contribution in data and sample collection. We also acknowledge the patients and families who contributed generously to the work that we performed.

REFERENCES

- Avraham, R., Haseley, N., Fan, A., Bloom-Ackermann, Z., Livny, J., and Hung, D. T. (2016). A highly multiplexed and sensitive RNA-seq protocol for simultaneous analysis of host and pathogen transcriptomes. *Nat. Protoc.* 11, 1477–1491. doi: 10.1038/nprot.2016.090
- Azoulay, E., Mokart, D., Kouatchet, A., Demoule, A., and Lemiale, V. (2019). Acute respiratory failure in immunocompromised adults. *Lancet Respir Med.* 7, 173–186. doi: 10.1016/S2213-2600(18)30345-X
- Azoulay, E., Russell, L., Van de Louw, A., Metaxa, V., Bauer, P., Pova, P., et al. (2020). Diagnosis of severe respiratory infections in immunocompromised patients. *Intensive Care Med.* 46, 298–314. doi: 10.1007/s00134-019-05906-5
- Barriere, S. L. (2010). Challenges in the design and conduct of clinical trials for hospital-acquired pneumonia and ventilator-associated pneumonia: an industry perspective. *Clin. Infect. Dis.* 51(Suppl. 1), S4–S9. doi: 10.1086/653033
- Cao, B., Huang, Y., She, D. Y., Cheng, Q. J., Fan, H., Tian, X. L., et al. (2018). Diagnosis and treatment of community-acquired pneumonia in adults: 2016 clinical practice guidelines by the Chinese thoracic society,

- chinese medical association. *Clin. Respir. J.* 12, 1320–1360. doi: 10.1111/crj.12674
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Chen, S., Zhu, Q., Xiao, Y., Wu, C., Jiang, Z., Liu, L., et al. (2021). Clinical and etiological analysis of co-infections and secondary infections in COVID-19 patients: an observational study. *Clin. Respir. J.* 15, 815–825. doi: 10.1111/crj.13369
- Chiu, C. Y., and Miller, S. A. (2019). Clinical metagenomics. *Nat. Rev. Genet.* 20, 341–355. doi: 10.1038/s41576-019-0113-7
- Di Pasquale, M. F., Sotgiu, G., Gramegna, A., Radovanovic, D., Terraneo, S., Reyes, L. F., et al. (2019). Prevalence and etiology of community-acquired pneumonia in immunocompromised patients. *Clin. Infect. Dis.* 68, 1482–1493. doi: 10.1093/cid/ciy723
- Dulanto Chiang, A., and Dekker, J. P. (2020). From the pipeline to the bedside: advances and challenges in clinical metagenomics. *J. Infect. Dis.* 221(Suppl._3), S331–S340. doi: 10.1093/infdis/jiz151
- Flygare, S., Simmon, K., Miller, C., Qiao, Y., Kennedy, B., Di Sera, T., et al. (2016). Taxonomer: an interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. *Genome Biol.* 17:111. doi: 10.1186/s13059-016-0969-1
- Forbes, J. D., Knox, N. C., Ronholm, J., Pagotto, F., and Reimer, A. (2017). Metagenomics: the next culture-independent game changer. *Front. Microbiol.* 8:1069. doi: 10.3389/fmicb.2017.01069
- Fournier, P. E., Dubourg, G., and Raoult, D. (2014). Clinical detection and characterization of bacterial pathogens in the genomics era. *Genome Med.* 6:114. doi: 10.1186/s13073-014-0114-2
- Goldberg, B., Sichtig, H., Geyer, C., Ledebner, N., and Weinstock, G. M. (2015). Making the leap from research laboratory to clinic: challenges and opportunities for next-generation sequencing in infectious disease diagnostics. *mBio* 6:e01888-15. doi: 10.1128/mBio.01888-15
- Gu, W., Miller, S., and Chiu, C. Y. (2019). Clinical metagenomic next-generation sequencing for pathogen detection. *Annu. Rev. Pathol.* 14, 319–338. doi: 10.1146/annurev-pathmechdis-012418-012751
- Han, D., Li, Z., Li, R., Tan, P., Zhang, R., and Li, J. (2019). mNGS in clinical microbiology laboratories: on the road to maturity. *Crit. Rev. Microbiol.* 45, 668–685. doi: 10.1080/1040841X.2019.1681933
- Hogan, C. A., Yang, S., Garner, O. B., Green, D. A., Gomez, C. A., Dien Bard, J., et al. (2020). 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. doi: 10.1093/cid/ciaa035
- Kalantar, K. L., Moazed, F., Christenson, S. C., Wilson, J., Deiss, T., Belzer, A., et al. (2019). Metagenomic comparison of tracheal aspirate and mini-bronchial alveolar lavage for assessment of respiratory microbiota. *Am. J. Physiol. Lung Cell Mol. Physiol.* 316, L578–L584. doi: 10.1152/ajplung.00476.2018
- Letourneau, A. R., Issa, N. C., and Baden, L. R. (2014). Pneumonia in the immunocompromised host. *Curr. Opin. Pulm Med.* 20, 272–279. doi: 10.1097/MCP.0000000000000051
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Miao, Q., Ma, Y., Wang, Q., Pan, J., Zhang, Y., Jin, W., et al. (2018). Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin. Infect. Dis.* 67(Suppl._2), S231–S240. doi: 10.1093/cid/ciy693
- Miller, S., Naccache, S. N., Samayoa, E., Messacar, K., Arevalo, S., Federman, S., et al. (2019). Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res.* 29, 831–842. doi: 10.1101/gr.238170.118
- Naccache, S. N., Federman, S., Veeraraghavan, N., Zaharia, M., Lee, D., Samayoa, E., et al. (2014). A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Res.* 24, 1180–1192. doi: 10.1101/gr.171934.113
- Peng, J. M., Du, B., Qin, H. Y., Wang, Q., and Shi, Y. (2021). Metagenomic next-generation sequencing for the diagnosis of suspected pneumonia in immunocompromised patients. *J. Infect.* 82, 22–27. doi: 10.1016/j.jinf.2021.01.029
- Prina, E., Ranzani, O. T., and Torres, A. (2015). Community-acquired pneumonia. *Lancet* 386, 1097–1108. doi: 10.1016/S0140-6736(15)60733-4
- Ramachandran, P. S., and Wilson, M. R. (2020). Metagenomics for neurological infections - expanding our imagination. *Nat. Rev. Neurol.* 16, 547–556. doi: 10.1038/s41582-020-0374-y
- Ramirez, J. A., Wiemken, T. L., Peyrani, P., Arnold, F. W., Kelley, R., Mattingly, W. A., et al. (2017). Adults hospitalized with pneumonia in the united states: incidence, epidemiology, and mortality. *Clin. Infect. Dis.* 65, 1806–1812. doi: 10.1093/cid/cix647
- Shi, Y., Huang, Y., Zhang, T. T., Cao, B., Wang, H., Zhuo, C., et al. (2019). Chinese guidelines for the diagnosis and treatment of hospital-acquired pneumonia and ventilator-associated pneumonia in adults (2018 Edition). *J. Thorac. Dis.* 11, 2581–2616. doi: 10.21037/jtd.2019.06.09
- Simner, P. J., Miller, S., and Carroll, K. C. (2018). Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin. Infect. Dis.* 66, 778–788. doi: 10.1093/cid/cix881
- Sousa, D., Justo, I., Dominguez, A., Manzur, A., Izquierdo, C., Ruiz, L., et al. (2013). Community-acquired pneumonia in immunocompromised older patients: incidence, causative organisms and outcome. *Clin. Microbiol. Infect.* 19, 187–192. doi: 10.1111/j.1469-0691.2012.03765.x
- Torres, A., Chalmers, J. D., Dela Cruz, C. S., Dominedò, C., Kollef, M., Martin-Loeches, I., et al. (2019). Challenges in severe community-acquired pneumonia: a point-of-view review. *Intens. Care Med.* 45, 159–171. doi: 10.1007/s00134-019-05519-y
- WHO (2020). *The Top 10 Causes of Death*. Geneva: WHO.
- Wilson, M. R., Sample, H. A., Zorn, K. C., Arevalo, S., Yu, G., Neuhaus, J., et al. (2019). Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N. Engl. J. Med.* 380, 2327–2340. doi: 10.1056/NEJMoa1803396
- Zhang, F., Chen, J., Huang, H., Deng, X., Zhang, W., Zeng, M., et al. (2021). Application of metagenomic next-generation sequencing in the diagnosis and treatment guidance of *Pneumocystis jirovecii* pneumonia in renal transplant recipients. *Eur. J. Clin. Microbiol. Infect. Dis.* 40, 1933–1942. doi: 10.1007/s10096-021-04254-x
- Zinter, M. S., Dvorak, C. C., Mayday, M. Y., Iwanaga, K., Ly, N. P., McGarry, M. E., et al. (2019a). Pulmonary metagenomic sequencing suggests missed infections in immunocompromised children. *Clin. Infect. Dis.* 68, 1847–1855. doi: 10.1093/cid/ciy802
- Zinter, M. S., Mayday, M. Y., Ryckman, K. K., Jelliffe-Pawlowski, L. L., and DeRisi, J. L. (2019b). Towards precision quantification of contamination in metagenomic sequencing experiments. *Microbiome* 7:62. doi: 10.1186/s40168-019-0678-6

Conflict of Interest: TX, FH, MX, and XL are employed by Vision Medicals Co., Ltd., Guangzhou, Guangdong Province, China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zhan, Xu, He, Guan, Li, Li, Xie, Li, Chen, Cheng, Zhong and Ye. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.