Original Article

Application of metagenomic next-generation sequencing (mNGS) in diagnosing pneumonia of adults

Zhiguang Liu¹, Chuang Sun¹, Xinru Xiao¹, Lianzheng Zhou¹, Yanhua Huang¹, Yujia Shi¹, Xiaowei Yin¹, Zhengdao Mao¹, Qian Zhang¹²

¹ Department of Respiratory and Critical Care Medicine, The Affiliated Changzhou Second People’s Hospital of Nanjing Medical University, Changzhou, China
² Changzhou Medical Center, Nanjing Medical University, Changzhou, China

Abstract

Introduction: Accurate identification of pathogens that cause pulmonary infections is essential for effective treatment and hastening recovery in adults diagnosed with pneumonia. At present, despite metagenomic next-generation sequencing (mNGS) technology has been widely used in clinical practice for pathogen identification, the clinical significance and necessity of detecting pathogen in bronchoalveolar lavage fluid (BALF) for pneumonia-stricken adults remain ambiguous.

Methodology: In this study, 80 patients suffering from pulmonary infection were enrolled, who were admitted to the Affiliated Changzhou Second People’s Hospital of Nanjing Medical University between January 2020 and September 2022. The diagnostic performances of mNGS and conventional methods (CM) were systematically analyzed based on BALF samples, and we further investigated the influence of mNGS and CM in diagnosis modification and treatment.

Results: We found a significantly higher positive rate for the mNGS method in contrast to CM. Bacteria were the most common pathogens, and Streptococcus pneumoniae was the most commonly identified pathogen. Candida albicans and Epstein-Barr virus were the most frequently identified fungus and virus. Atypical pathogens such as Mycobacterium tuberculosis, virus Nontuberculous mycobacteria, and Chlamydia psittaci were also identified. A total of 77 patients were identified with mixed infections by mNGS. As the disease progressed and recurrent antibiotic treatment persisted, significant dynamic changes in the clinical manifestation from the BALF samples could be found by mNGS.

Conclusions: This study underscores the efficacy of mNGS in detecting pathogens in BALF samples from patients suffering pulmonary infections. Compared with the CM, mNGS significantly enhanced the positive diagnosis ratio, particularly in diagnosing Mycobacterium tuberculosis, atypical pathogens, and viral or fungal infections.

Key words: Metagenomic next-generation sequencing; conventional method; pulmonary infection; pathogen.


(Received 06 June 2023 – Accepted 30 August 2023)

Copyright © 2023 Liu et al. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Pulmonary infection has the highest morbidity rates, and it is the most common infectious disease, causing a severe disease burden to society and individuals [1,2]. Diverse pathogens causing pulmonary infection leads to the continuous emergence of the original infectious pathogens and new pathogens. Therefore, the identification of pathogens is still the focus of clinicians. Conventional methods (CM) for the identification of pathogens include culture, microscopy, time of flight mass spectrometer (TOFMS), and immunology and polymerase chain reaction (PCR) [3-5], which have exposed several weaknesses in the diagnosis of mixed infections, such as long culture cycles and the influence of antibiotics limit the accuracy of diagnosis. Most importantly, CM has a narrow detection range and can detect suspicious pathogens but cannot diagnose unknown and new pathogens, which may lead to delayed or missed diagnosis [6]. Failure to diagnose and treat infection timely and accurately, may contribute to continued transmission of the pathogen and increased mortality among hospitalized patients [7]. Hence, it is necessary to apply new, superior methods for pathogen detection and diagnosis confirmation to guide targeted treatment and improve prognoses.

mNGS is a method for parallel sequencing all nucleic acids in a sample. Because mNGS can detect rare, new, difficult-to-detect, and co-infection pathogens, it has a particular application value in diagnosing. mNGS not only has a broad detection spectrum, including bacteria, viruses, fungi, atypical
pathogens, parasites, and even new microorganisms but it is also barely affected by antibiotics in diagnosed efficiency [8]. Although many studies have shown the diagnosis value of mNGS in suspected pulmonary, several problems need to be addressed, such as differentiation of colonization from infection, method standardization, and biological information standardization [9,10]. As a result, the diagnostic efficiency value of mNGS in pulmonary infection still needs to be explored.

In this retrospective clinical study, we ultimately enrolled 80 patients to investigate the diagnostic performances of mNGS and CM, and the pathogen distribution of pulmonary infection, and to investigate the clinical application value of mNGS in pulmonary infection.

**Methodology**

**Study patients**

Patients with suspected pneumonia were selected from the Department of Respiratory and Critical Care Medicine of Changzhou Second People's Hospital affiliated to Nanjing Medical University, recruited from January 2020 to September 2022, a total of 153 patients were investigated by reviewing electronic medical records. Patients with suspected pneumonia that met the following criteria were enrolled: (i) there were related manifestations in imaging, such as pulmonary exudation; (ii) clinical manifestations included fever, cough, expectoration, respiratory failure, and other typical manifestations of pulmonary infection. The hospital ethics committee approved this study.

Final patient inclusion criteria were as follows: (i) Patients with pneumonia who consent to undergo bronchoscopy to collect BALF, and both mNGS and CM were performed to detect pathogens; (ii) the quality inspection and BALF sample testing process met the standards of mNGS (The total DNA of BLAF was ≥ 10 ng and the concentration was ≥ 0.3 ng/µL); (iii) the patient clinical data were complete; (iv) the raw mNGS sequence data were complete. Excluded were patients who had not been diagnosed with pneumonia and could not meet the above requirements from enrollment. Pneumonia was diagnosed based on a composite reference standard derived from the diagnostic criteria of community-acquired pneumonia (CAP) [11] and hospital-acquired pneumonia (HAP) [12]. Figure 1 shows the design of this study.

**Sample preparation**

BALF samples were collected for mNGS and CM detection in patients with pulmonary infection according to standard procedures. The patients were routinely fasted for 8 hours before surgery and received 0.5 mg atropine intramuscular injection and 2% lidocaine inhalation under local anesthesia. A fibrous bronchoscope was introduced through the nasal cavity to observe the state of the trachea, bronchi, and lobular bronchi at the lesion site, and the tip of the fibrous bronchoscope was inserted into the bronchi at the infected lesion site and embedded in the lumen. During lavage, about 50 mL of 37 ℃ saline water was injected through the biopsy hole and repeated about 3 times until the lavage solution was clarified and the lavage solution was sucked out. The BALF samples were sent to the laboratory for immediate analysis upon collection at room temperature. A total of 20 mL BALF samples were divided into aliquots and about 5 mL of BALF was used for nucleic acid extraction for mNGS detection.

In addition, sputum and blood samples from some patients were also collected for CM. Samples of 3 mL sputum from each patient were collected in sterile tubes and were liquefied using 0.1% dithiothreitol for 30 minutes at room temperature.

**Conventional methods**

Conventional testing included microbial culture, smears: acid-fast staining for *Mycobacterium*, PCR assays, and serological antibody detection for *Chlamydia*, *Streptococcus pneumoniae*, *Epstein-Barr virus*, *Cytomegalovirus*, and other herpes simplex viruses. Additionally, 1,3-β-D-glucan and galactomannan were performed for *Candida* and *Aspergillus*, respectively. T-spot and Xpert testing were done for *Mycobacterium tuberculosis*.

**mNGS Procedure for BALF Samples**

The total DNA of the BALF sample was extracted using the Tiangen Magnetic DNA Kit (Tiangen). DNA
libraries were constructed through DNA fragmentation, end-repair, add A-tailing, adapter ligation, and PCR amplification and were performed using the NEB Next® Ultra™ DNA Library Prep Kit for Illumina®. The Agilent 2100 bioanalyzer detected the quality of the library. The concentration of the library was detected by Qubit 2.0. The qualified libraries were sequenced using the Illumina Next-seq platform. The raw data was split using bcl2fastq2, and the adaptor sequences and low-quality base sequences were removed using Trimmomatic software to obtain high-quality, effective data. Human host sequences were removed using bowtie2 calibration software. Finally, sequences not mapped to the human genome were retained for alignment with the microbial genome database. The pathogenic microorganism database is constructed by screening standard microbial nucleic acid sequences in public databases. To compare the abundance of different species within the same sample, the number of reads was normalized by the length of the species genome to calculate their “Reads Per Kilobase (RPK).” The relative abundance of the species was further calculated based on RPK.

Results Description
The detection list was obtained after filtering through the threshold criterion, and the thresholds were set as the number of reads stringently mapped to the species of bacterium, mycoplasma, chlamydia, DNA virus or fungus ≥ 3, and Mycobacterium tuberculosis complex (MTC) ≥ 1. Finally, two or more associate chief physicians determined the final diagnosis about the causative agents by referring to the content of the sequencing reports, such as the number of unique reads, relative abundance, information of sequenced samples, and the microbial genome size, comprehensively querying microbial pathogenicity, and combining with the clinical characteristics and examination results of the patients.

Statistical Analysis
Continuous variables were described by medians, and categorical variables were expressed as counts and percentages. We tested for the differences in categorical variables with a chi-squared test. Data analyses were performed using GraphPad Prism 7 software and a two-tailed \( p \) value of 0.05 was considered to represent a significant difference.

Results
General Characteristics of the Patients
A total of 153 patients with suspected pneumonia were reviewed in this retrospective cohort study between January 2020 to September 2022, and among them, 96 patients underwent bronchoscopy, and BALF samples were collected for pathogen detection by mNGS and conventional methods. Sixteen patients (16/96, 16.67%) were diagnosed with non-infectious pneumonia. Of the remaining 80 patients, there were 55 males and 25 females, and 14 patients (14/80, 17.50%) were admitted to the Respiratory Intensive Care Unit (RICU) and received mechanical ventilation. The average age of these enrolled adults was 65 years old, ranging from 32 to 97 years old. According to the Chinese Guidelines for Management of Community-Acquired Pneumonia in Adults (2016) [13], a total of 62 patients had underlying diseases, including hypertension, diabetes, tumor, hepatitis, coronary heart disease, chronic obstructive pulmonary disease (COPD), cerebral infarction, bronchitis, bronchiectasis, anemia, chronic gastritis, epilepsy, hyperthyroidism, pulmonary tuberculosis. Multiple underlying conditions in many adults were common. Evaluated the outcome of antibiotic treatment in each patient based on the clinical symptoms (included fever, shortness of breath, respiratory rate, lung rales, and clinical manifestations), imaging, white blood cell count, C-reactive protein level, procalcitonin level, oxygen saturation, and other laboratory and routine detection indicators. The baseline characteristics of patients are shown in Table 1.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Clinical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (avg)</td>
<td>65 (32-97)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>55 (55/80, 68.75%)</td>
</tr>
<tr>
<td>Female</td>
<td>25 (25/80, 31.25%)</td>
</tr>
<tr>
<td>Diagnostic results</td>
<td></td>
</tr>
<tr>
<td>Severe pneumonia</td>
<td>14 (14/80, 17.50%)</td>
</tr>
<tr>
<td>Non-severe pneumonia</td>
<td>66 (66/80, 82.50%)</td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>Underlying diseases</td>
<td>62 (62/80, 77.50%)</td>
</tr>
<tr>
<td>Without underlying diseases</td>
<td>18 (18/80, 22.50%)</td>
</tr>
<tr>
<td>Admitted to the RICU</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>14 (14/80, 17.50%)</td>
</tr>
<tr>
<td>No</td>
<td>66 (66/80, 82.50%)</td>
</tr>
<tr>
<td>Mechanical ventilation</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>14 (14/80, 17.50%)</td>
</tr>
<tr>
<td>No</td>
<td>66 (66/80, 82.50%)</td>
</tr>
<tr>
<td>Antibiotic exposure before mNGS</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>77 (77/80, 96.25%)</td>
</tr>
<tr>
<td>No</td>
<td>3 (3/80, 3.75%)</td>
</tr>
<tr>
<td>Hospital stays (days)</td>
<td></td>
</tr>
<tr>
<td>≥ 14</td>
<td>7 (7/80, 8.75%)</td>
</tr>
<tr>
<td>&lt; 14</td>
<td>73 (73/80, 91.25%)</td>
</tr>
</tbody>
</table>
Pathogens’ profile of all pneumonia patients according to mNGS and conventional methods

A total of 189 strains of pathogens were identified in 80 cases by a combination of mNGS and CM detection (Figure 2). Bacteria were the most common pathogens as there were 116 (116/189, 61.38%) strains. The most frequently detectable Gram-negative bacteria were *Haemophilus parainfluenzae*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, and *Pseudomonas aeruginosa*. The most common Gram-positive bacteria were *Streptococcus pneumoniae*, *Enterococcus faecium*, and *Staphylococcus aureus*. A total of 30 fungi were detected in 34 cases, and the most common were *Candida albicans* and *Aspergillus fumigatus*. A total of 26 viruses were identified, including *Epstein-Barr virus*, *Cytomegalovirus*, *Human herpesvirus 1*, and *Human betaherpesvirus 7*. A total of 3 Mycobacteria were detected, which included *Mycobacterium tuberculosis*, *Mycobacterium Abscessus*, and *Mycobacterium avium complex*; additionally, 8 cases were confirmed as non-typical pathogens such as *Chlamydia psittaci* and *Tropheryma whippelii*.

Comparison of diagnostic performance between mNGS and conventional methods

The results of mNGS and CM were both positive in 42 (42/80, 52.50%) cases, a total of 38 (38/80, 47.50%) cases were positive by mNGS only. In addition, for 42 cases, the results between mNGS and CM were entirely consistent in 5 (5/80, 6.25%), partially consistent in 35 (35/80, 43.75%), and utterly inconsistent in 2 (2/80, 2.50%) (Figure 3). mNGS had a higher bacterial positive rate than CM (100% vs. 68.75%, \( p < 0.0001 \)).

Some difficult-cultured bacteria, including nontuberculosis mycobacteria (n = 4), *Tropheryma whippelii* (n = 3), *Chlamydia psittaci* (n = 4), and Nocardia (n = 2), were only detected by mNGS. For fungi, except *Candida albicans*, other fungi such as *Pneumocystis jirovecii* (n = 1), *Aspergillus flavus* (n = 3), and *Aspergillus fumigatus* (n = 6) were only detected by mNGS. Furthermore, mNGS detected viruses in 12 cases, while conventional detection methods were negative for almost all viruses. Ultimately, through symptoms, clinical manifestation, imaging examinations, CM and mNGS comprehensively diagnosed 66 (66/80, 82.5%) patients.

Single infection and mixed infection

There were 80 patients confirmed pulmonary infection. A total of 3 patients were confirmed to have a single infection, and 77 patients were diagnosed with a mixed infection with two or more pathogens. In all mixed infection cases, the most common pathogens were bacteria-bacteria (46/77, 59.7%), bacteria-fungi (30/77, 39.0%), and bacteria-viruses (1/77, 1.30%). 3 patients with a single infection were identified as infected with bacteria (Figure 4).

Effect of mNGS on treatment and prognosis

The initial empirical treatment completely covered the pathogens in 37.5% (30/80) of the patients. Then, the empirical treatment strategy was adjusted based on microbiological results, laboratory examination results, and response to the initial treatment. A total of 50 patients received adjustments: 10 patients were adjusted based on clinical experience, and 40 patients were adjusted based on microbiological results. 6 patients were adjusted according to CM results only, 29 patients were adjusted according to mNGS results only, and 5 patients were adjusted based on CM combined mNGS results. Among the cases adjusted based solely on mNGS, antifungal drugs were initiated in 14 cases,
antiviral medicines in 2 cases, and anti-tuberculosis drugs in 6 cases. mNGS contributed to the adjustment of 34 patients, with a higher improvement rate than patients adjusted solely based on CM results (76.47% vs. 66.67%, \(p = 0.6091\)). In total, 67 (67/80, 83.75%) patients were modified, 10 (10/80, 12.50%) patients showed disease progression, 2 (2/80, 2.50%) patients were transferred to another hospital for treatment, and one patient (1/80, 1.25%) ultimately passed away.

**Discussion**

The diversity of patients and symptoms, multiple infections, and the advent of new pathogens are progressing complexities to the clinical treatment of anti-pneumonia infection. In response to these issues, we screened 80 patients in whom mNGS and CM tests were used. Our two-year comprehensive analysis and epidemiological supplement aimed to evaluate the mNGS diagnostic performance with regard to pneumonia infection in this region. We selected BALF samples for pathogen detection mainly because they could better reflect the actual state of the lung compared with sputum and throat swabs and had higher detection sensitivity and specificity [14,15]. In a prospective study, 329 adult patients obtained an equal volume of BALF to detect pathogens by mNGS and CM with severe community-acquired pneumonia; it was identified in 304 cases (99.24%) in the mNGS group, while that of CM was 39.5% [16]. We also used mNGS to detect BAFL samples, showing a very high microbial detection rate (80/80, 100%), significantly outperforming CM. Similarly, other articles related to pulmonary infection also reported that the positive rate of mNGS was higher than that of CM’s [17,18].

As a direct method for nucleic acid microorganism detection, mNGS offers a broader pathogen detection spectrum compared to CM. There were some atypical pathogens of pulmonary infection, including *Chlamydia psittaci*, *Nocardia asiatica*, *Mycoplasma*, which had also been reported by mNGS testing [19-21]. This study identified some pathogens detected almost exclusively by mNGS, such as *Mycobacterium tuberculosis* and Nocardia, which may require longer latency, as well as some non-culturable pathogens under standard conditions, such as *Chlamydia psittaci*, non-tuberculous mycobacteria. The viruses are always ignored by Clinicians, because they frequently colonize the respiratory tract and are relatively harmless. In contrast, the viruses have recently been identified as severe end-stage disease instigators in patients with immune dysfunction or immunosuppression [22-24]. mNGS solved the current clinical difficulties in detecting the virus and was the primary method to find viruses, especially newly emerging viruses that cause pneumonia, such as SARS-COV [25,26]. In this study, all viruses were detected by the mNGS technique, including the notably threatening cytomegalovirus to immunocompromised patients [27]. mNGS also has superior feasibility in detecting of fungi, especially for difficult cultivation fungi and cases with a low fungal load. Between 2014 and 2021, over 300 cases of fungal infection diagnosed by NGS were reported, and *Pneumocystis jirovecii* was the predominant fungus reported (25%), followed by *Aspergillus* (22%) and *Candida* (16%) [28]. It should be noted that *Aspergillus* and *Pneumocystis jirovecii* were only detected by mNGS in this study, which complemented CM and substantially improved the overall fungal detection rate. By the way, our findings additionally highlighted a surge in *Candida albicans* detection. This increase in results was attributed to the high false-positive rate of mNGS and CM due to contamination. *Candida albicans* were widely distributed in the environment, consequently, the quality of the samples seriously affected testing and analysis outcomes, urging clinicians to follow sterile sampling principles [29]. It is crucial to note that nearly all patients in this study were empirically treated with antibiotics before sampling, which was responsible for reducing the sensitivity of CM. Nonetheless, mNGS remained unaffected, probably due to the extended survival time of pathogenic DNA.

Regarding the diagnosis of mixed infection, the proportion of mixed lung infections in our patients was high, and most of them were mixed bacterial-bacterial, which was consistent with the results of a previous study [30]. Multiple studies have reported that mNGS was significantly superior to conventional tests, such as

![Figure 4. Distribution of Single infection or Mixed infection by mNGS and CM.](image-url)
serological antibodies and cultures, in identifying co-infections [31,32]. Various microorganisms interact and inhibit each other, making identification difficult by traditional culture methods [33]. However, it was known that mixed infections usually caused severe pneumonia. Therefore, mNGS was a more favorable method for diagnosing pneumonia patients.

mNGS technology can guide clinical changes in antibiotic treatment, thereby enhancing patient prognosis. In our study, 62.5% of the patients depended on clinical features, the results of mNGS and CM, and other examinations to adjust the final diagnosis and treatment. Of these, the antibiotic regimen of 34 patients was adjusted based on mNGS results, which positively impacted the course and prognosis (Table 2). Especially crucial, benefit from the high sensitivity of mNGS in detecting fungi, viruses, and Mycobacterium tuberculosis, we initiated the targeted antibiotic therapy. With the widespread clinical application of mNGS, clinicians have begun to adjust the initial empirical antibiotic use according to mNGS detection. Benefitting from the rapid and accurate diagnosis of mNGS, early targeted antibiotics could significantly improve the prognosis of patients with respiratory system infections [16,34].

Although mNGS is commonly used in clinical practice, the clinical application of mNGS remains challenging due to the lack of unified standards. There standards include indications of mNGS, appropriate sampling times, quality control, sequencing platforms, data analysis, and interpretation of results. In addition, it is difficult to explain the results of mNGS, the main problem is distinguishing between colonization and infection, especially opportunistic infection. A study has shown the use of bioinformatics to differentiate between colonization and infection, but this study was necessary due to the small sample [35]. Additionally, prohibitive cost of mNGS testing restricts its accessibility to many families, which limited the choice of mNGS first and finally delayed diagnosis and treatment. We believe that reducing the cost of mNGS will benefit patients.

In conclusion, we analyzed and compared the detection results from mNGS and CM. Since it was a small retrospective study, we were unable to conclude the correlation between mNGS and clinical indicators. Nevertheless, the results indicate that mNGS can improve the diagnosis of pulmonary infections, particularly Mycobacterium tuberculosis, atypical pathogens, and viral and fungal infections. Further, mNGS shows great promise in diagnosing mixed infections, adjusting antibiotic use, and improving patient prognosis.

Acknowledgements
The authors would like to thank the Genoxor Medical Science and Technology Inc to complete the whole process of mNGS and serious data collection.

Authors’ contributions
All authors made a significant contribution to the work. Qian Zhang and Zhengdao Mao designed the study. Zhiguang Liu collected and analyzed the data. Zhiguang Liu and Chuang Sun drafted the paper. All authors revised the paper.

Funding
This study was supported by grants from the Jiangsu Province social development project (BE2020651 to Q.Z.), and in part from "333 Project" of Jiangsu Province (BRA2020015 to Q.Z.), the Changzhou Sci & Tech Program (CE20205023 to Q.Z.), and the Changzhou High-Level Medical Talents Training Project (2022CZLJ013 to Q.Z.), China Postdoctoral Science Foundation (2020M670011ZX to Zhengdao Mao).

Ethical Approval and Consent to participate
The present study followed the Declaration of Helsinki and was approved by the ethics committee of Changzhou Second People’s Hospital of Nanjing Medical University ([2019]KY032-01). Written informed consent was obtained from all patients or their guardians at the beginning of the study.

<table>
<thead>
<tr>
<th>All patients (n = 40)</th>
<th>Antibiotic use according to mNGS result only</th>
<th>Antibiotic use according to CM result only</th>
<th>Antibiotic use according to CM and mNGS result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Anti-fungal</td>
<td>Anti-viral</td>
</tr>
<tr>
<td>Positive prognosis</td>
<td>29</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>Negative prognosis</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


Corresponding authors
Qian Zhang, PhD
Changzhou Medical Center,
Nanjing Medical University, Changzhou, 213003, China
Tel: +86-0519-81087391
Email: qianzhang@njmu.edu.cn

Zhengdao Mao, MD
Department of Respiratory and Critical Care Medicine,
The Affiliated Changzhou Second People’s Hospital of Nanjing Medical University,
Changzhou, 213003, China
Tel: +86-13915012305
Email: mzd_czey@njmu.edu.cn

Conflict of interests: No conflict of interests is declared.