The diagnostic value of metagenomic next-generation sequencing in infectious diseases

CURRENT STATUS: UNDER REVIEW

BMC Infectious Diseases  ▫  BMC Series

Hongxia Duan  
Shanghai Tenth People's Hospital

Xuan Li  
Shanghai Tenth People's Hospital

aihong mei  
Shanghai Tenth People's Hospital

Ping Li  
Shanghai Tenth People's Hospital

Yang Liu  
Shanghai Tenth People's Hospital

Xiaofeng Li  
Shanghai Tenth People's Hospital

Weiwei Li  
Shanghai Tenth People's Hospital

Shuanshuan Xie  
Shanghai Tenth People's Hospital

Changhui Wang  
Department of Respiratory Medicine, Shanghai Tenth people's Hospital, Tongji University, Shanghai 200072, China

w@wang-chang-hui@hotmail.com  
Corresponding Author

ORCID: https://orcid.org/0000-0001-6118-6126

DOI: 10.21203/rs.3.rs-18934/v1

SUBJECT AREAS
   Infectious Diseases
KEYWORDS

Next-generation sequencing, Sensitivity, Diagnostic, Infection, Survival
Abstract

**Background and Objective:** Although traditional diagnostic techniques of infection are mature and price favorable at present, most of them are time-consuming and with a low positivity. Metagenomic next-generation sequencing (mNGS) was studied widely because of identification and typing of all pathogens not rely on culture and retrieving all DNA without bias. Based on this background, we aim to detect the difference between mNGS and traditional culture method, and to explore the relationship between mNGS results and the severity, prognosis of infectious patients.

**Methods:** 109 patients were enrolled in our study in Shanghai Tenth People’s Hospital from October 2018 to December 2019. The diagnostic results, negative predictive values, positive predictive values, false positive rate, false negative rate, pathogen and sample types were analyzed by using both traditional culture and mNGS methods. Then, the samples and clinical information of 93 patients in the infected group (ID) were collected. According to whether mNGS detected pathogens, the patients in ID group were divided into the positive group of 67 cases and the negative group of 26 cases. Peripheral blood leukocytes, C-reactive protein (CRP), procalcitonin (PCT) and neutrophil counts were measured, and the concentrations of cytokines in the serum were determined by ELISA. The correlation between the positive detection of pathogens by mNGS and the severity of illness, hospitalization days, and mortality were analyzed.

**Results:** 109 samples were assigned into infected group (ID, 92/109, 84.4%), non-infected group (NID, 16/109, 14.7%), and unknown group (1/109, 0.9%). Blood was the most abundant type of samples with 37 cases, followed by alveolar lavage fluid (BALF) in 36 cases, tissue, sputum, pleural effusion, pus and so on. In the ID group, the majority of patients were diagnosed with lower respiratory system infections (73/109, 67%), followed by bloodstream infections, pleural effusion and central nervous system infections. The sensitivity of mNGS was significantly higher than that of culture method (67.4% vs 23.6%; \( P < 0.001 \)), especially in sample tapes of BALF (\( P=0.002 \)), blood (\( P<0.001 \)) and sputum (\( P=0.037 \)), while the specificity of mNGS was not significantly different from culture method (68.8% vs 81.3%; \( P = 0.41 \)). The number of hospitals stays and 28-day-mortality in the positive mNGS group were significantly higher than those in the negative group, and the difference
was statistically significant (P<0.05). Age was significant in multivariate logistic analyses of positive results of mNGS.

**Conclusions:** The study found that mNGS had a higher sensitivity than the traditional method, especially in blood, BALF and sputum samples. And positive mNGS group had a higher hospital stays 28-day-mortality, which means the detection of pathogen nucleic acid sequences may be a potential high-risk factor for poor prognosis of patients and has significant clinical value. MNGS should be used more in early pathogen diagnosis in the future.

**Introduction**
As is well-known, infectious diseases are still the leading cause of morbidity and mortality worldwide and spread quickly. As the first-line department of pathogen detection, microbiology laboratory plays an important role in infection control by means of microscopic examination, culture, identification, drug sensitivity and so on. However, the limitation of molecular diagnosis and genotyping methods remain pathogens undetected in up to 60% of cases. Failure to identify pathogens in time may delay the precise treatment of antibiotics, leading to unnecessary use of broad-spectrum antibiotics, inducing resistance, and increasing medical costs.

With the completion of the human genome project in the early 21st century and the rapid development of sequencing technology, high-throughput and low-cost second-generation sequencing technology emerged. It had been used in whole genome sequencing, whole exome sequencing, macro gene sequencing and so on, among which metagenomic next-generation sequencing (mNGS) was studied most widely. The advantage of mNGS lies in the single run to obtain the sequence information of microbial nucleic acid fragments, through analysis and comparison of which to detect all microbial species and sequence. Besides, mNGS can be used for the identification and typing of all pathogens because mNGS does not rely on culture and retrieve all DNA without bias. Therefore, this technology may play a huge role in infection prevention and medical microbiology laboratory. Thus, based on microbiome sequencing technology, we compared the sensitivity and specificity of mNGS method and traditional culture method to detect pathogens, and discussed the influence of
mNGS detection results on the severity and prognosis of patients with infection in our study.

Methods
Study Patients
We retrospectively reviewed 161 cases suspected of acute or chronic infection at Shanghai 10th People’s Hospital in Shanghai, China, between October 2018 and December 2019. Excluding pregnancy, mental illness and patients under the age of 18, 109 samples were included for analysis and categorized into 3 groups defined as infectious disease (ID), noninfectious disease (NID), and unknown groups according to final diagnosis. Specimens were subjected to regular clinical microbiological assay as well as mNGS testing (BGI, Intertek, Biotecan, China) in a pairwise manner. Meanwhile, clinical data of all enrolled patients, including blood routine, CRP, PCT, neutrophil count and cytokines were collected. This research had been approved by the ethics committee of the 10th People's Hospital affiliated to Tongji University.

Metagenomic Next-generation Sequencing and Analysis
Nucleic acid detection and sequencing were performed based on BGISEQ-50 platform in this research. After the sample was taken, nucleic acid was extracted, the library was built and sequenced, and finally the data was analyzed by using the microbiome database. The experimental process was shown in Fig. 1.

Sample Processing and Library construction (Fig. 1A)
For infected patients or patients with fever of unknown cause, infected site samples were collected according to standard procedures. Since most of the collected samples contain pathogenic pathogens, they were inactivated before nucleic acid extraction. In addition to this, blood samples were centrifuged to separate plasma and leukocytes according to the actual situation and sputum samples were liquefied by using 0.1% DTT (dithiothreitol) for 30 minutes at room temperature after inactivation. After that, DNA were extracted by TIANamp Micro DNA Kit (DP316, Tiangen Biotech) according to the manufacturer’s recommendation, and the nucleic acid quality is the key factor of mNGS detection. DNA libraries were constructed by end-repair method. Each trial included internal, negative and positive controls. Internal parameters helped identify failed or abnormal samples to control the quality of DNA.
Bioinformatic analysis (Fig. 1B)

1. Quality control
Effective sequencing data quantity was not less than 20 M. a. Sequencing subtracted of human host sequences need to be above 90%; b. Reads of microbial detection sequences were longer than 50 bp for the homology of microbiology sequence is high. Considering the interference of human genome and sensitivity of detection, it is recommended that the effective sequencing data volume should not be less than 20M without removing the human genome component.

2. Data filtering
In order to obtain high quality sequence data, the qualified data was further filtered by bioinformatics analysis to remove low quality sequences. High quality sequences were then removed human host sequences. Currently, the most commonly referenced host sequences are human genome reference sequences. Different comparison methods designed different algorithms according to different comparison strategies.

3. Sequences alignment
The filtered sequences were compared with the reference sequences in the pathogen database, which covers bacteria, fungi, viruses, protozoa and other pathogenic microorganisms. For successful alignment sequences, it is necessary to further remove high coverage repeat sequences and low-quality alignment sequences, and to exclude multiple alignment as much as possible, so as to obtain the final pathogen alignment results. According to the final results of pathogen comparison, all parameters of detected pathogens were calculated, including sequence number, relative abundance, genome coverage and depth, etc.

4. Report generation
In the process of report generation, the suspected background microbial bank and negative control test results were referred to, and combined with the clinical information provided, the final test reports were given.

Determination of cytokines
Detection of cytokines (TNF-a, IL-2, IL-4, IL-6, IL-8, IL-10, IL-17A and INF-r) in serum was by solid phase, sandwich and chemiluminescence using the IMMULITE/IMMULIE 1000 analyzer. The chemiluminescence kit was from SIEMENS, Germany. The processed specimens were sent to the analyzer for testing according to the manufacturer's instructions, and the corresponding cytokine concentrations were recorded.

Cell classification and count detection
Cells were classified using the American Thermo Fisher automatic flow cytometer and divided into total white blood cells, neutrophil count, CD4 + T cell count, CD8 + T cell count, B cells, and NK, T cell count.

Statistical Analysis
Comparative analysis was conducted by Pearson $\chi^2$ test and t test. Data analyses was performed by using SPSS 22.0 software. P values < 0.05 were considered significant, and all tests were 2-tailed. Logistic regression analysis explored the risk factors associated with positive detection of mNGS.

Results
Sample and Patient Characteristics
Demographic features of the patients were provided in Table 1. 87 males and 22 females participated in our study, whose average age was 61 years old, average length of stay was 17.53 days and the case fatality rate were 11.93%. Most (37/109, 33.9%) of our samples were from blood, 36 of 109 (33.0%) were from BALF, 12 of 109 (11.0%) were from tissue and 9 (8.3%) of 109 were from sputum, followed by pleural fluid (7, 6.4%), CSF(4, 3.7%), pus (2, 1.8%), bone marrow (1, 0.9%) and swab (1, 0.9%) (Fig. 2A). In the study cohort, 92 (84.4%) patients diagnosed with confirmed pathogens by the conventional technique were assigned to ID group. The remaining specimens were subdivided into the NID (16/109, 14.7%) and unknown (1/109, 0.9%) groups (Fig. 2B). There were no statistical differences between ID group and NID group in age, gender, length of stay and case fatality rate ($p > 0.05$ in all). The majority of patients were diagnosed with respiratory system infections (73/109, 67.0%), followed by bloodstream infections (10/109, 9.17%), pleural effusion (6/109, 5.50%) and central nervous system infections (6/109, 5.50%) as shown in Fig. 2C.
Table 1
Demographic characteristic of samples

|                  | Total  | ID      | NID     | Unknown | P value between ID & NID |
|------------------|--------|---------|---------|---------|-------------------------|
| Samples amount, n (%) | 109(100%) | 92(84.40) | 16(14.68) | 1(0.92) | /                       |
| Age, average years (range) | 61.02(25-95) | 60.26(25-95) | 66(40-90) | 61(/) | 0.43                     |
| Gender, male, n (%) | 87(79.82) | 74(80.43) | 12(75.00) | 1(100%) | 0.62                     |
| Length of stay, average days (range) | 17.53(1-70) | 16.88(1-70) | 20.87(6-61) | 22(/) | 0.31                     |
| Case fatality rate, % | 11.93 | 13.04 | 6.25 | 0 | 0.39                     |

Diagnostic Performance Comparison of mNGS and Culture

Comparison of Diagnostic Performance for Differentiating ID From NID

The cases of mNGS and culture tests for the ID, NID, and unknown groups were illustrated in Fig. 3A.

In the chi-square test of positive rate, there were statistical differences between mNGS and culture of all and of ID group, but no differences in NID and unknown group for the limited amounts. To compare the diagnostic efficiency for differentiating ID from NID, 105 samples were included for further study.

The negative predictive values and positive predictive values of diagnosing infectious disease by mNGS were 27.5% and 92.3%, respectively, with the negative likelihood ratio and positive likelihood ratio being 0.47 and 2.16. As expected, mNGS increased the sensitivity rate by approximately 44% in comparison with that of culture (67.4% vs 23.6%; P < 0.001) and decreased the specificity rate by 12.5% compared with that of culture (68.8% vs 81.3%; P = 0.41) (Fig. 3B).

Concordance Between mNGS and Culture for Pathogen Detection

In our results, mNGS and culture were both positive in 21 of 109 (19.3%) cases and were both negative in 25 of 109 (22.9%) cases. 58 samples were positive by mNGS only (53.2%) and 5 were positive by culture only (4.6%). For double-positive samples, the 2 results were completely matched in 3 of 21 cases and totally mismatched in 3 of 21 cases (Fig. 3C). The remaining 15 cases were found to be “partly matched,” indicative of at least one overlap of pathogens when polymicrobial results were observed.

“False Positives” and “False Negatives” of mNGS

In the ID group, up to three culturable pathogens were missed by mNGS. Among these “mNGS false-negative” cases, 2 culture results were paradoxical with clinical diagnosis, and the remaining 1 were
completely unidentifiable by mNGS. For the 7 “mNGS false-positive” cases in the NID group, possible reasons included potential concomitant infection with NIDs (3/7), overinterpretation (3/7) and unknown (1/7) (Table 2).

| Sample No. | Specimen source | Diagnosis | mNGS result | Possible explanation |
|------------|----------------|-----------|-------------|----------------------|
| 2          | BALF           | Hematencephalon | Acinetobacter baumannii, Klebsiella, Enterococcus | Unknown |
| 33         | Blood          | Lymphoma   | Pseudomonas, CMV | Potential cause of lymphoma |
| 62         | Blood          | Aplastic anemia | Acinetobacter baumannii, Enterococcus | Overinterpretation |
| 67         | Blood          | myelofibrosis | Phycomyces blakesleeanus | Overinterpretation |
| 74         | Pleural Fluid  | Pleural effusion | Fusobacterium nucleatum, Streptococcus constellatus, Porphyromonas gingivalis | Potential cause of inflammation |
| 86         | Blood          | Ulcerative Colitis | Porphyromonas gingivalis, HSV | Potential cause of inflammation |
| 88         | Blood          | Lung cancer  | Saccharomyces cerevisiae | Overinterpretation |

Culturable Pathogens Missed by mNGS in the ID Group

| Microbe | Count | Possible explanation |
|---------|-------|----------------------|
| MTB     | 2     | Positive Not Detected |
| Pseudomonas | 1 | Microbes “Weak” |

Abbreviations: mNGS, metagenomic next-generation sequencing; ID, infectious disease; NID, noninfectious disease; CMV, metagenomic next-generation sequencing; HSV, herpes simplex virus; MTB, Mycobacterium tuberculosis.

Comparison Analysis at the Pathogen-type Level

Among the 69 microbes isolated, Klebsiella (10/69) was the most commonly detected pathogen, followed by bacteria without MTB/NTM (9/69), Aspergillus (6/69), Pseudomonas (6/69) and EBV (6/69) (Fig. 4A). The percentage of mNGS-positive samples observed to have a higher yield rate by mNGS than that by culture, but the differences were not significant (P > 0.05) in terms of Klebsiella, bacteria without MTB/NTM, EBV, CMV due to the small sample size. In Acinetobacter baumannii (n = 2) and MTB (n = 3), the number of mNGS-positive samples was equally with that of culture-positive samples. While only mNGS indicated positive results in NTM (n = 4), Anaerobes (n = 4), Saccharomyces cerevisiae (n = 2), Proteus (n = 1), Pneumocystis carinii (n = 2), Abiotrophia (n = 1), Nocardia (n = 3), Staphylococcus aureus (n = 2), Enterococcus (n = 2) and Escherichia coli (n = 1).

Comparison Analysis at the Sample-type Level
In the types of BALF, blood and sputum samples, mNGS detection had significantly higher sensitivity than the culture method (P = 0.002 for BALF, P < 0.001 for blood, P = 0.037 for sputum), and the overall sensitivity of mNGS in the sample types was significantly different (P = 0.03). In addition, in the culture method, the positive rate in BALF was higher than that in the whole blood (P = 0.019), and there was no difference in the overall sensitivity of the culture method in the sample type, as shown in Fig. 4B.

Comparison of infection indexes in positive and negative group by mNGS in ID Classification and counting of leukocyte and lymphocyte in positive and negative group by mNGS

In this study, routine blood tests, CRP and PCT tests were examined on the day of examination of pathogenic microorganisms to determine the differences in the total number of white blood cells, lymphocytes and neutrophils between the positive group and the negative group by mNGS. The results showed (Table 3) that there were no statistically differences in leukocyte and lymphocyte (P > 0.05).

| Cytokines pg/ml | Positive    | Negative    | P       |
|-----------------|-------------|-------------|---------|
| IL-2            | 100.35 ± 68.21 | 1.31 ± 0.94 | 0.511   |
| IL-4            | 2.74 ± 0.41  | 1.52 ± 0.94 | 0.206   |
| IL-6            | 70.8 ± 18.27 | 68.96 ± 33.18| 0.964   |
| TNF-α           | 2.48 ± 0.42  | 2.26 ± 1.32 | 0.842   |
| IL-17a          | 13.77 ± 2.35 | 10.45 ± 8.01| 0.592   |
| IL-8            | 1154 ± 0     | -           | -       |
| IL-10           | 26.14 ± 7.75 | 8.29 ± 3.33 | 0.044   |
| IFN-γ           | 8.91 ± 1.89  | 13.59 ± 6.92| 0.361   |
| Cellular Immunity % |         |             |         |
| CD4/CD8         | 1.42 ± 0.23  | 2.06 ± 0.44 | 0.185   |
| Th cell         | 35 ± 3.36    | 43.83 ± 5.75| 0.201   |
| Ts cell         | 68.06 ± 3.07 | 66.67 ± 3.64| 0.18    |
| NK cell         | 15.71 ± 2.17 | 15.5 ± 1.89 | 0.958   |
| B cell          | 13.53 ± 2.94 | 12.83 ± 3.73| 0.899   |
| T cell          | 68.06 ± 3.07 | 66.67 ± 3.64| 0.958   |
| WBC×10^9        | 8.32 ± 0.52  | 7.36 ± 0.48 | 0.283   |
| Neu×10^9        | 6.99 ± 0.58  | 5.38 ± 0.48 | 0.109   |
| PCT ng/ml       | 0.34 ± 0.17  | 3.42 ± 3.32 | 0.112   |
| CRP mg/l        | 87.63 ± 8.32 | 63.61 ± 13.47| 0.129   |

Abbreviations: mNGS, metagenomic next-generation sequencing; WBC, white blood cells; IL-, interleukin-; IFN-γ, Interferon-γ; TNF-α, Tumor Necrosis Factor-α; CD4, Cluster of Differentiation 4 receptors; CD8, Cluster of Differentiation 8 receptors; Th, helper T cell; Ts, suppressor T cell; NK, natural killer cell; Neu, neutrophil; PCT, procalcitonin; CRP, C-reactive protein.

Comparison of cytokine concentrations in positive and negative group by mNGS

In order to explore the correlation between the status of immune function in patients and the positive
results of pathogen examination, this study detected and analyzed the peripheral blood (TNF-α, IL-2, IL-4, IL-6, IL-8, IL-10, IL-17A and INF-γ) in infected patients. The results indicated that the peripheral blood concentrations of IL-10 in the positive group was higher than that in the negative group, and the differences were statistically significant (P = 0.044), while other cytokine showed no difference between groups as shown in Table 3.

Analysis of correlative factors for positive result of pathogen extraction by mNGS

In order to further explore the related risk factors of positive mNGS test in infected patients, this study used Logistic multivariate regression analysis to analyze the collected patients’ information and whether the pathogen was detected in the patients. After the confounding factors were removed, the variables that were significant for detection was age (P = 0.037, OR:1.076, 95% CI:1.005–1.152), which promoted the detection of pathogens (Table 4).

| Values        | B   | SE (B) | Wald X² | P     | OR    | 95% CI    |
|---------------|-----|--------|---------|-------|-------|-----------|
| Age           | 0.073 | 0.035 | 4.367   | 0.037 | 1.076 | 1.005-1.152 |
| Sex           | -0.545 | 1.157 | 0.222   | 0.637 | 0.58  | 0.06-5.601  |
| Read Number   | -2.371 | 0.599 | 15.677  | 0.000 | 0.093 | 0.029-0.302 |
| HOD           | -0.028 | 0.061 | 0.216   | 0.642 | 0.972 | 0.863-1.095 |
| Survival Time | -0.007 | 0.005 | 1.888   | 0.169 | 0.993 | 0.983-1.003 |
| IL-2          | 0.171 | 1.115 | 0.023   | 0.878 | 1.186 | 0.133-10.553 |
| IL-4          | -1.299 | 0.893 | 2.116   | 0.146 | 0.273 | 0.047-1.57  |
| IL-6          | -0.005 | 0.019 | 0.077   | 0.781 | 0.995 | 0.957-1.033 |
| TNF-α         | -0.374 | 0.373 | 1.003   | 0.316 | 0.688 | 0.331-1.430 |
| IL-17α        | 0.202 | 0.137 | 2.165   | 0.141 | 1.223 | 0.935-1.6   |
| IL-10         | -2.64 | 0.206 | 1.639   | 0.201 | 0.768 | 0.513-1.151 |
| IFN-γ         | 0.09  | 0.071 | 1.606   | 0.205 | 1.095 | 0.952-1.259 |
| Cytokines pg/ml|     |       |         |       |       |           |
| IL-2          | 0.171 | 1.115 | 0.023   | 0.878 | 1.186 | 0.133-10.553 |
| IL-4          | -1.299 | 0.893 | 2.116   | 0.146 | 0.273 | 0.047-1.57  |
| IL-6          | -0.005 | 0.019 | 0.077   | 0.781 | 0.995 | 0.957-1.033 |
| TNF-α         | -0.374 | 0.373 | 1.003   | 0.316 | 0.688 | 0.331-1.430 |
| IL-17α        | 0.202 | 0.137 | 2.165   | 0.141 | 1.223 | 0.935-1.6   |
| IL-10         | -2.64 | 0.206 | 1.639   | 0.201 | 0.768 | 0.513-1.151 |
| IFN-γ         | 0.09  | 0.071 | 1.606   | 0.205 | 1.095 | 0.952-1.259 |
| Cell Immunity %|     |       |         |       |       |           |
| CD4/CD8       | -0.488 | 0.965 | 0.256   | 0.613 | 0.614 | 0.093-4.067 |
| Th cell       | 0.318 | 0.296 | 1.511   | 0.283 | 1.374 | 0.769-2.454 |
| Ts cell       | 0.244 | 0.317 | 0.589   | 0.443 | 1.276 | 0.685-2.377 |
| NK cell       | -0.223 | 0.211 | 1.121   | 0.29  | 0.800 | 0.529-1.209 |
| B cell        | -0.265| 0.245 | 1.172   | 0.279 | 0.767 | 0.227-1.475 |
| T cell        | 0.5485 | 0.478 | 1.315   | 0.252 | 0.578 | 0.475-1.239 |
| WBC×10⁹       | -0.123 | 1.228 | 0.01    | 0.92  | 0.884 | 0.08-9.819  |
| Neu×10⁹       | 0.141 | 1.39  | 0.01    | 0.919 | 1.151 | 0.076-17.535 |
| PCT ng/ml     | -0.681 | 1.514 | 0.202   | 0.653 | 0.506 | 0.026-9.844 |
| CRP mg/l      | -0.004 | 0.015 | 0.073   | 0.788 | 0.996 | 0.968-1.025 |

Abbreviations: HOD, hospital day; WBC, white blood cells; IL-, interleukin-; IFN-γ, Interferon-γ; TNF-α, Tumor Necrosis Factor-α; CD4, Cluster of Differentiation 4 receptors; CD8, Cluster of Differentiation 8 receptors; Th, helper T cell; Ts, suppressor T cell; NK, natural killer cell; Neu, neutrophil; PCT, procalcitonin; CRP, C-reactive protein.

Potential Implications of Clinical mNGS Test

Potential Inappropriate Antibiotic Usage for Patients with Virus Isolates

Among the 4 viruses from 23 patients, the most commonly identified viruses were herpes simplex
virus (n = 15), followed by Epstein-Barr virus/ herpes simplex virus (n = 5), Epstein-Barr virus (n = 1), Hepatitis A virus (n = 1) and torque teno virus (n = 1). Almost one-half of patients were diagnosed with a hospital-acquired infection (12/23) and there was a considerable percentage of patients (17/23) prescribed broad-spectrum antibiotics, while 10 of 23 patients were suspected of inappropriate antibiotic usage and 7 of 23 were considered immunocompromised hosts (Table 5).

| Type of Virus | HAI | Immunosuppressed Patients | Broad-spectrum Antibiotics | Suspected Inappropriate Antibiotic Usage | Treatment Responsive |
|---------------|-----|---------------------------|----------------------------|-----------------------------------------|---------------------|
| HSV (n = 15)  | Yes | 8                         | Yes                        | Yes                                     | Yes                 |
|               | No  | 7                         | No                         | Yes                                     | Yes                 |
| HAV (n = 1)   | Yes | 0                         | No                         | Yes                                     | Yes                 |
|               | No  | 1                         | Yes                        | No                                      | Yes                 |
| HSV/EBV (n = 5) | Yes | 3                         | Yes                        | Yes                                     | Yes                 |
|               | No  | 2                         | No                         | Yes                                     | Yes                 |
| TTV (n = 1)   | Yes | 0                         | Yes                        | Yes                                     | Yes                 |
|               | No  | 1                         | No                         | Yes                                     | Yes                 |
| Total (N = 23)| Yes | 12                        | 17                         | 10                                      | 12                  |
|               | No  | 11                        | 6                          | 10                                      | 11                  |

The influence of positive by mNGS on the hospital days and survival of patients

As Table 6 showed, there were 67 samples in positive group with 57 males and 26 in negative group with 20 males. There was no significant difference in mean age between the two groups (59.70 yrs vs 60.50 yrs, P = 0.84). Positive group had a longer hospital day (HOD, 176.63 days vs 150.96 days, P = 0.047) and a higher 28-day mortality (9.0% vs 0%, P = 0.049) than those of negative group, but there were no statistical differences in 14-day mortality (4.5% vs 0%, P = 0.278) and 90-day mortality (13.4% vs 3.9%, P = 0.180) between groups. The average survival time of two groups were 176.64 days and 150.96 days, respectively, but P value for t test between groups was 0.425, no statistical differences. The survival curves of the two groups were shown in Fig. 5. At the meantime, we analyzed the relationship between pathogens read number and HOD, 14-day-mortality, 28-day-mortality and 90-day-mortality, which showed that the higher pathogens read number, the higher 90-day-mortality and the longer HOD (Table 7).
Table 6
The basic demographic and clinical characteristics of initial and outcome patient variables.

|                      | Positive    | Negative    | P    |
|----------------------|-------------|-------------|------|
| Sex                  |             |             |      |
| Female               | 10          | 6           | 0.355|
| Male                 | 57          | 20          |      |
| Age                  | 59.70 ± 2.16| 60.50 ± 3.06| 0.84 |
| HOD                  | 176.63 ± 17.70| 150.96 ± 103.14| 0.047|
| 14 days of death     | 4.5%        | 0           | 0.278|
| 28 days of death     | 9.0%        | 0           | 0.049|
| 90 days of death     | 13.4%       | 3.9%        | 0.180|
| Read Number          | 5295.62 ± 2507.26| 16.67 ± 4.79| 0.039|
| Survival time        | 176.64 ± 17.70| 150.96 ± 21.05| 0.425|

Abbreviations: HOD, hospital day.

Table 7
The analysis between the pathogens read number and HOD, 14, 28 and 90-day-mortality

| Read Number | F   | P    |
|-------------|-----|------|
| No          | 20  |      |
| 1–9         | 15  | 24   |
| 10–99       | 24  |      |
| 100–999     | 14  |      |
| 1000–9997   |     |      |
| HOD         | 14.84 ± 8.58| 13.07 ± 5.18| 15.80 ± 9.12| 20.70 ± 16.5| 27.92 ± 24.06| 2.685| 0.037|
| 14-mortality| 0.05| 0.04| 0.29|
| 28-mortality| 0.05| 0.04| 0.29|
| 90-mortality| 0.05| 0.29| 0.36|
| Survival    | 169.74 ± 102.68| 138.40 ± 100.27| 158.70 ± 125.83| 185.45 ± 124.82| 194.71 ± 216.79| 0.424| 0.791|

Abbreviations: HOD, hospital day.

Discussion
The traditional clinical model for diagnosing infectious diseases is for doctors to make a differential diagnosis and then conduct a series of tests to try to identify the pathogen. Traditional diagnostic techniques of microbiology laboratory ranges from smear microscopy, microorganisms’ culture, antigen antibody detection and PCR mainly. Whereas most traditional methods were often time-consuming and has a lower positive rate. Although molecular diagnostic assays are a quick way to diagnose the most common infections, almost all conventional microbial trials in use today only target a limited number of pathogens at a time or require successful culture of microorganisms from clinical samples. While mNGS analyse the entire microbiome in patients’ samples so it has been used to discover novel viral pathogens and diagnose viral infections in people widely. Therefore, we explored the application and differences between traditional culture method and mNGS in clinical infectious diseases. BALF, blood, sputum, tissue, CSF, pleural fluid, pus, bone marrow or swab from 109 patients suspected of infection were collected and specimens were subjected to regular clinical microbiological assay and mNGS testing in a pairwise manner in our study. We then systematically
compared the clinical features and test results of mNGS and traditional culture. The results suggested that there were no significant differences in age, gender, length of stay and fatality rate between two groups and mNGS had an advantage in sensitivity with no significant inferior in specificity rate compared with traditional culture method. A team of researchers also found that mNGS detected potential pathogenic bacteria, which had advantages in speed and sensitivity compared with culture and pathology, Miao’s team showed that mNGS had a sensitivity and specificity of 50.7% and 85.7% for the diagnosis of infectious diseases, higher than traditional culture (50.7% vs 35.2%). In particular, the diagnosis of MTB, virus, anaerobic bacteria, nocardia and fungi has obvious advantages. The results were similar to our results, which showed that the sensitivity of mNGS was 67.4% significantly higher than that of culture method (23.6%). High sensitivity of mNGS may because pathogen DNA has a long survival time in plasma, the use of antibiotics has a small impact on mNGS results, while traditional cultures are greatly affected by the use of antibiotics. Because of the small sample size, mNGS showed no statistical difference compared with culture method in pathogen types although there was a trend of superiority in Klebsiella, bacteria without MTB/NTM, EBV, CMV, NTM, Anaerobes, Saccharomyces cerevisiae, Proteus, Pneumocystis carinii, Abiotrophia, Nocardia, Staphylococcus aureus, Enterococcus and Escherichia coli. However, mNGS detection had a significantly higher sensitivity than the culture method in BALF (P = 0.002), blood (P < 0.001) and sputum (P = 0.037) samples.

Based on the advantages shown by mNGS, we then investigated the influence of positive mNGS detection results on the severity and prognosis of patients with infection. By comparing the classification and counting of leukocyte, lymphocyte and cytokine concentrations in positive and negative groups, we found that IL-10 concentration in peripheral blood in the positive group was higher than that in the negative group and there were no statistically differences in another cytokine concentration, leukocyte and lymphocyte. According to the results of correlative factors analysis for positive test of mNGS, patients’ age may promote the detection of pathogens. In the survival analysis, positive group had a higher 28-day mortality (9.0% vs 0%, P = 0.049) than that of negative group, but there were no statistical differences in average survival time. And the pathogens read number by
mNGS was positive related to the HOD and 90-day-mortality of patients with infectious diseases. All of that indicated older people were more likely to have positive results and positive results of mNGS detection may represent a worse outcome.

Fortunately, mNGS has moved from scientific application to clinical practice and is changing the way disease diagnosed and treated\textsuperscript{19–21}. In addition to what we mentioned above, mNGS also has merits in many other aspects. Firstly, mNGS does not need prior clinical information to detect infectious pathogens, and the results can be reported quickly and accurately, greatly shortening the diagnosis time of infectious pathogens. As we all knows, early and rapid reporting of the results by mNGS provides clinical clues to the next step in diagnosis and treatment, especially avoiding overuse of antibiotics for viral infections\textsuperscript{22, 23}. Rapid results reported by mNGS also can promote timely adjustment of treatment in clinical practice. As our data showed, almost one-half of patients were suspected of inappropriate antibiotic usage. Secondly, mNGS was used in some rare infectious pathogens. It detected naegleria fowleri\textsuperscript{24}, brucellosis\textsuperscript{25}, cysticercosis, taenia bocinea\textsuperscript{26}, gondii\textsuperscript{27} in CSF, Hepatic tuberculosis in blood\textsuperscript{28} in previous reports. Thirdly, studies have shown that mNGS can be used not only for pathogen identification, but also for microbiome characterization, parallel analyses of human host responses, drug resistance gene and virulence factor detection. All of these led to the rapid development of mNGS in immunodeficiency difficult-to-diagnose cases and immunocompromised patients\textsuperscript{14}. However, mNGS still has some limitations at present, such as human background, background bacteria contamination, no uniform standards for detailed experimental procedures\textsuperscript{3, 29–32}, inability to distinguish infection and colonization, standardization of bioinformatics analysis process, and problem of report interpretation. The results must be interpreted in the context of the clinical situation. It’s worth noting that background microbial contamination is a common problem faced by mNGS technology, which can be partially eliminated through negative quality control, but it requires clinical familiarity with common background bacteria and better interpretation results combined with clinical practice\textsuperscript{25}.

In this study, we systematically compared mNGS and traditional culture method in sensitivity,
specificity, pathogen type and sample type. On this basis, we also compared and analyzed the differences between the positive and negative groups of mNGS which was few at present. Patients of positive group found to have a trend of worse prognosis suggested need more attention clinically. Small sample size was the biggest deficiency of our study, so that there were many results indicated a certain trend without reaching statistical significance unfortunately. Therefore, more patients need to be included in the study in the future. Not randomized controled was also the limitation of study. In summary, mNGS had a higher sensitivity, especially in the types of BALF, blood and sputum samples, and there was a trend of higher sensitivity of Klebsiella, CMV and EBV detection. The worse trend of outcome in patients with positive mNGS results than negative group prompted more clinical attention is required. Therefore, based on what we found above and other advantages of mNGS like quick results, less affected by prior antibiotic exposure and so on, we suggest that mNGS should be used more in early pathogen diagnosis in the future. Nonetheless, interpreting data of mNGS will be a challenge for doctors in guiding clinical treatment of infectious diseases.

Declerations

Acknowledgments

We would like to thank all the patients who donate their biological samples.

Author Contributions

Conception and design: H.X.D., S.S.X., X.L., P.L. Acquisition, statistical analysis or interpretation of the data: all authors. Drafting of the manuscript: C.H.W. All authors reviewed and approved the final version of the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (No. 81802262), the Fundamental Research Funds for the Central Universities (No. 22120180584), Shanghai Tenth Hospital's Improvement Plan for NSFC (No. 04.03.17.032, 04.01.18.048, SYGZRPY2017014).

Compliance with ethical standards

Conflict of interest

The authors declare no potential conflicts of interest related to this study.
References

1 Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Agarwal R, Ahn SY, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Barker-Collo S, Bartels DH, Bell ML, Benjamin EJ, Bennett D, Bhalla K, Bikbov B, Bin Abdulhak A, Birbeck G, Blyth F, Bolliger I, Boufous S, Bucello C, Burch M, Burney P, Carapetis J, Chen H, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahodwala N, De Leo D, Degenhardt L, Delossantos A, Denenberg J, Des Jarlais DC, Dharmaratne SD, Dorsey ER, Driscoll T, Duber H, Ebel B, Erwin PJ, Espindola P, Ezzati M, Feigin V, Flaxman AD, Forouzanfar MH, Fowkes FG, Franklin R, Fransen M, Freeman MK, Gabriel SE, Gakidou E, Gaspari F, Gillum RF, Gonzalez-Medina D, Halasa YA, Haring D, Harrison JE, Havmoeller R, Hay RJ, Hoen B, Hotez PJ, Hoy D, Jacobsen KH, James SL, Jasrasaria R, Jayaraman S, Johns N, Karthikeyan G, Kassebaum N, Keren A, Khoo JP, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lipnick M, Lipshultz SE, Ohno SL, Mabweijano J, MacIntyre MF, Mallinger L, March L, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGrath J, Mensah GA, Merriman TR, Michaud C, Miller M, Miller TR, Mock C, Mocumbi AO, Mokdad AA, Moran A, Mulholland K, Nair MN, Naldi L, Narayan KM, Nasseri K, Norman P, O'Donnell M, Omer SB, Ortblad K, Osborne R, Ozgediz D, Pahari B, Pandian JD, Rivero AP, Padilla RP, Perez-Ruiz F, Perico N, Phillips D, Pierce K, Pope CA, 3rd, Porrini E, Pourmalek F, Raju M, Ranganathan D, Rehm JT, Rein DB, Remuzzi G, Rivara FP, Roberts T, De Leon FR, Rosenfeld LC, Rushton L, Sacco RL, Salomon JA, Sampson U, Sanman E, Schwebel DC, Segui-Gomez M, Shepard DS, Singh D, Singleton J, Sliwa K, Smith E, Steer A, Taylor JA, Thomas B, Tleyjeh IM, Towbin JA, Truelsen T, Undurraga EA, Venketasubramanian N, Vijayakumar L, Vos T, Wagner GR, Wang M, Wang W, Watt K, Weinstock MA, Weintraub R, Wilkinson JD, Woolf AD, Wulf S, Yeh PH, Yip P, Zabetian A, Zheng ZJ, Lopez AD, Murray CJ, AlMazroa MA, Memish ZA. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012; 380:2095-2128.

2 Zhou K, Lokate M, Deurenberg RH, Tepper M, Arends JP, Raangs EG, Lo-Ten-Foe J, Grundmann H,
Rossen JW, Friedrich AW. Use of whole-genome sequencing to trace, control and characterize the regional expansion of extended-spectrum beta-lactamase producing ST15 Klebsiella pneumoniae. Sci Rep 2016; 6:20840.

3 Schlaberg R, Chiu CY, Miller S, Procop GW, Weinstock G, Professional Practice C, Committee on Laboratory Practices of the American Society for M, Microbiology Resource Committee of the College of American P. Validation of Metagenomic Next-Generation Sequencing Tests for Universal Pathogen Detection. Arch Pathol Lab Med 2017; 141:776-786.

4 Ewig S, Torres A, Angeles Marcos M, Angrill J, Rano A, de Roux A, Mensa J, Martinez JA, de la Bellacasa JP, Bauer T. Factors associated with unknown aetiology in patients with community-acquired pneumonia. Eur Respir J 2002; 20:1254-1262.

5 van Gageldonk-Lafeber AB, Heijnen ML, Bartelds AI, Peters MF, van der Plas SM, Wilbrink B. A case-control study of acute respiratory tract infection in general practice patients in The Netherlands. Clin Infect Dis 2005; 41:490-497.

6 Miao Q, Ma Y, Wang Q, Pan J, Zhang Y, Jin W, Yao Y, Su Y, Huang Y, Wang M, Li B, Li H, Zhou C, Li C, Ye M, Xu X, Li Y, Hu B. Microbiological Diagnostic Performance of Metagenomic Next-generation Sequencing When Applied to Clinical Practice. Clin Infect Dis 2018; 67:S231-S240.

7 Grumaz S, Stevens P, Grumaz C, Decker SO, Weigand MA, Hofer S, Brenner T, von Haeseler A, Sohn K. Next-generation sequencing diagnostics of bacteremia in septic patients. Genome Med 2016; 8:73.

8 Lecuit M, Eloit M. The diagnosis of infectious diseases by whole genome next generation sequencing: a new era is opening. Front Cell Infect Microbiol 2014; 4:25.

9 Lefterova MI, Suarez CJ, Banaei N, Pinsky BA. Next-Generation Sequencing for Infectious Disease Diagnosis and Management: A Report of the Association for Molecular Pathology. J Mol Diagn 2015; 17:623-634.

10 Khare R, Espy MJ, Cebelinski E, Boxrud D, Sloan LM, Cunningham SA, Pritt BS, Patel R, Binnicker MJ. Comparative evaluation of two commercial multiplex panels for detection of gastrointestinal pathogens by use of clinical stool specimens. J Clin Microbiol 2014; 52:3667-3673.

11 Leber AL, Everhart K, Balada-Llasat JM, Cullison J, Daly J, Holt S, Lephart P, Salimnia H,
Schreckenberger PC, Desjarlais S, Reed SL, Chapin KC, LeBlanc L, Johnson JK, Soliven NL, Carroll KC, Miller JA, Dien Bard J, Mestas J, Bankowski M, Enomoto T, Hemmert AC, Bourzac KM. Multicenter Evaluation of BioFire FilmArray Meningitis/Encephalitis Panel for Detection of Bacteria, Viruses, and Yeast in Cerebrospinal Fluid Specimens. J Clin Microbiol 2016; 54:2251-2261.

12 Ruggiero P, McMillen T, Tang YW, Babady NE. Evaluation of the BioFire FilmArray respiratory panel and the GenMark eSensor respiratory viral panel on lower respiratory tract specimens. J Clin Microbiol 2014; 52:288-290.

13 Tang YW, Gonsalves S, Sun JY, Stiles J, Gilhuley KA, Mikhлина A, Dunbar SA, Babady NE, Zhang H. Clinical Evaluation of the Luminex NxTAG Respiratory Pathogen Panel. J Clin Microbiol 2016; 54:1912-1914.

14 Chiu CY, Miller SA. Clinical metagenomics. Nat Rev Genet 2019; 20:341-355.

15 Moustafa A, Xie C, Kirkness E, Biggs W, Wong E, Turpaz Y, Bloom K, Delwart E, Nelson KE, Venter JC, Telenti A. The blood DNA virome in 8,000 humans. PLoS Pathog 2017; 13:e1006292.

16 Rascovan N, Duraisamy R, Desnues C. Metagenomics and the Human Virome in Asymptomatic Individuals. Annu Rev Microbiol 2016; 70:125-141.

17 Somasekar S, Lee D, Rule J, Naccache SN, Stone M, Busch MP, Sanders C, Lee WM, Chiu CY. Viral Surveillance in Serum Samples From Patients With Acute Liver Failure By Metagenomic Next-Generation Sequencing. Clin Infect Dis 2017; 65:1477-1485.

18 Li H, Gao H, Meng H, Wang Q, Li S, Chen H, Li Y, Wang H. Detection of Pulmonary Infectious Pathogens From Lung Biopsy Tissues by Metagenomic Next-Generation Sequencing. Front Cell Infect Microbiol 2018; 8:205.

19 Houldcroft CJ, Beale MA, Breuer J. Clinical and biological insights from viral genome sequencing. Nat Rev Microbiol 2017; 15:183-192.

20 Schlaberg R, Queen K, Simmon K, Tardif K, Stockmann C, Flygare S, Kennedy B, Voelkerding K, Bramley A, Zhang J, Eilbeck K, Yandell M, Jain S, Pavia AT, Tong S, Ampofo K. Viral Pathogen Detection by Metagenomics and Pan-Viral Group Polymerase Chain Reaction in Children With Pneumonia Lacking Identifiable Etiology. J Infect Dis 2017; 215:1407-1415.
21 Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, Salamat SM, Somasekar S, Federman S, Miller S, Sokolic R, Garabedian E, Candotti F, Buckley RH, Reed KD, Meyer TL, Seroogy CM, Galloway R, Henderson SL, Gern JE, DeRisi JL, Chiu CY. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. N Engl J Med 2014; 370:2408-2417.

22 Rhodes J, Hyder JA, Peruski LF, Fisher C, Jorakate P, Kaewpan A, Dejsirilert S, Thamthitiwat S, Olsen SJ, Dowell SF, Chantra S, Tanwisaid K, Maloney SA, Baggett HC. Antibiotic use in Thailand: quantifying impact on blood culture yield and estimates of pneumococcal bacteremia incidence. Am J Trop Med Hyg 2010; 83:301-306.

23 Gosiewski T, Ludwig-Galezowska AH, Huminska K, Sroka-Oleksiak A, Radkowski P, Salamon D, Wojciechowicz J, Kus-Slowinska M, Bulanda M, Wolkow PP. Comprehensive detection and identification of bacterial DNA in the blood of patients with sepsis and healthy volunteers using next-generation sequencing method - the observation of DNAemia. Eur J Clin Microbiol Infect Dis 2017; 36:329-336.

24 Guo LY, Feng WY, Guo X, Liu B, Liu G, Dong J. The advantages of next-generation sequencing technology in the detection of different sources of abscess. J Infect 2019; 78:75-86.

25 Fan S, Ren H, Wei Y, Mao C, Ma Z, Zhang L, Wang L, Ge Y, Li T, Cui L, Wu H, Guan H. Next-generation sequencing of the cerebrospinal fluid in the diagnosis of neurobrucellosis. Int J Infect Dis 2018; 67:20-24.

26 Hu Z, Weng X, Xu C, Lin Y, Cheng C, Wei H, Chen W. Metagenomic next-generation sequencing as a diagnostic tool for toxoplasmic encephalitis. Ann Clin Microbiol Antimicrob 2018; 17:45.

27 Ai JW, Li Y, Cheng Q, Cui P, Wu HL, Xu B, Zhang WH. Diagnosis of local hepatic tuberculosis through next-generation sequencing: Smarter, faster and better. Clin Res Hepatol Gastroenterol 2018; 42:178-181.

28 Du B, Tao Y, Ma J, Weng X, Gong Y, Lin Y, Shen N, Mo X, Cao Q. Identification of sparganosis based on next-generation sequencing. Infect Genet Evol 2018; 66:256-261.

29 Blauwkamp TA, Thair S, Rosen MJ, Blair L, Lindner MS, Vilfan ID, Kawli T, Christians FC, Venkatasesubrahmanyam S, Wall GD, Cheung A, Rogers ZN, Meshulam-Simon G, Huijse L, Balakrishnan S, Quinn JV, Hollemon D, Hong DK, Vaughn ML, Kertesz M, Bercovici S, Wilber JC, Yang S. Analytical
and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. Nat Microbiol 2019; 4:663-674.

30 Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, Garcia-Cobos S, Kooistra-Smid AM, Raangs EC, Rosema S, Veloo AC, Zhou K, Friedrich AW, Rossen JW. Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol 2017; 243:16-24.

31 Gargis AS, Kalman L, Lubin IM. Assuring the Quality of Next-Generation Sequencing in Clinical Microbiology and Public Health Laboratories. J Clin Microbiol 2016; 54:2857-2865.

32 Miller S, Naccache SN, Samayoa E, Messacar K, Arevalo S, Federman S, Stryke D, Pham E, Fung B, Bolosky WJ, Ingebrigtsen D, Lorizio W, Paff SM, Leake JA, Pesano R, DeBiasi R, Dominguez S, Chiu CY. Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. Genome Res 2019; 29:831-842.

Abbreviations

mNGS, metagenomic next-generation sequencing; BALF, bronchoalveolar lavage Fluid; CSF, cerebrospinal fluid; Dx, diagnosis; ID, infectious disease; NID, noninfectious disease; EBV, Epstein-Barr virus; HAI, hospital-acquired infection; HboV, human bocavirus; HSV-1, herpes simplex virus 1; TTV, torquetenovirus; HAV, Hepatitis A virus; HOD, hospital day; IL-2, Interleukin-2; IL-4, Interleukin-4; IL-6, Interleukin-6; IL-10, Interleukin-10; IL-8, Interleukin-10; CD4, Cluster of Differentiation 4 receptors; CD8, Cluster of Differentiation 8 receptors; WBC, white blood cell; Th, helper T; NK, natural killer cell; PCT, procalcitonin; CRP, C-reactive protein; IFN-γ, Interferon-γ; IL-17a, Interleukin-17a; TNF-α, Tumor Necrosis Factor-α; PPV, positive predictive value; NPV, negative predictive value; CMV, cytomegalovirus; MTB, Mycobacterium tuberculosis; ns, no significant difference; NTM, nontuberculous mycobacteria; G+, Gram Positive; G-, Gram Negative.

Figures
Figure 1

Flow diagram of Metagenomic Next-generation Sequencing and Analysis.
Patients composition and samples types. A. In samples of this study, 33.9% were from blood which was the most, 33.0% from BALF, 11.0% from tissue and the others were from sputum (8.3%), pleural fluid (6.4%), CSF (3.7%), pus(1.8%), bone marrow(0.9%) and swab (0.9%). B. Patients were subdivided into ID (92/109, 84.4%), NID (16/109, 14.7%) and unknown (1/109, 0.9%) groups according to their diagnosis by conventional technique. C. Infection sites of patients in ID group. Most were respiratory system infections (73/109, 67.0%) and followed by bloodstream infections (10/109, 9.17%), pleural effusion (6/109, 5.50%), central nervous system infections (6/109, 5.50%), cardiovascular system infection (2/109,1.83%), eye, ear, nose, throat, or mouth infection (2/109,1.83%), skin and soft tissue infection (1/109, 0.92%), multifocal infection(1/109, 0.92%), urinary system infection(1/109, 0.92%). Abbreviations:

CSF, cerebrospinal fluid.
Diagnostic Performance Comparison of mNGS and Culture. A. Positive and negative cases in all, ID, NID and unknown group of mNGS and the culture, respectively. There were statistical differences between mNGS and culture of all (P<0.01) and of ID group (P<0.01), but no differences in NID and unknown group for the limited amounts (P>0.05). B. Contingency tables showed the sensitivity and specificity of mNGS were 67.4% and 68.8%, while those of culture were 23.6% and 81.3%. mNGS increased the sensitivity in comparison with that of culture (P < 0.001) while there were no differences in specificity between them (P = 0.41). C. Pie chart demonstrated the positivity distribution of mNGS and culture for all samples from 3 groups. 53.21% were positive by mNGS, 4.59% by culture, 19.27% by both and 22.94% were both negative. Abbreviations: NPV, negative predictive values; PPV, positive predictive values.
The overlap of positivity between mNGS and culture in pathogen and sample types. A. 19 pathogens detected in ID group with their corresponding frequencies were showed in histograms. Klebsiella, bacteria without MTB/NTM, EBV, CMV, NTM, Anaerobes, Saccharomyces cerevisiae, Proteus, Pneumocystis carinii, Abiotrophia, Nocardia, Staphylococcus aureus, Enterococcus and Escherichia coli demonstrated a trend of higher positivity rate in mNGS than that in culture with no statistical differences (P>0.05).

Acinetobacter baumannii and MTB were found equally in two groups. B. The overall sensitivity of mNGS in the different sample types were significantly different (P=0.03) while sample types did not affect the sensitivity of pathogens in culture. Interestingly, especially in the types of BALF, blood and sputum samples, mNGS had significantly higher sensitivity than the culture (P=0.002 for BALF, P<0.001 for blood, P=0.037 for sputum). Abbreviations: BALF, bronchoalveolar lavage fluid; CSF, cerebrospinal fluid; mNGS, metagenomic next-
The survival curves of positive and negative group of mNGS in ID. The survival curves suggested that the overall survival rate declined faster in the positive group, however, there was no statistically differences between the two groups.