Applying the pathogen-targeted next-generation sequencing method to pathogen identification in cerebrospinal fluid

Daiquan Gao¹[^], Yongqiang Hu²[^], Xuebin Jiang³[^], Hao Pu⁴[^], Zhendong Guo⁴[^], Yunzhou Zhang¹[^]

¹Department of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, China; ²Department of Critical Care Medicine, Beijing Fengtai You'anmen Hospital, Beijing, China; ³Intensive Care Unit, Renhe Hospital, Beijing, China; ⁴Department of Science and Technology, Shanghai Pathogeno Medical Technology Co., Ltd., Shanghai, China

Contributions: (I) Conception and design: D Gao; (II) Administrative support: Y Zhang; (III) Provision of study materials or patients: D Gao, Y Hu; (IV) Collection and assembly of data: X Jiang; (V) Data analysis and interpretation: H Pu, Z Guo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yunzhou Zhang, MD, PhD. Department of Neurology, Xuanwu Hospital, Capital Medical University, No. 45 Changchun Street, Xicheng District, Beijing 100053, China. Email: yuz657@sina.com.

Background: The cerebrospinal fluid (CSF) culture is a widely used method for the diagnosis of meningitis, but its detection sensitivity is low. Several new methods have been developed for pathogen detection, including metagenomic next-generation sequencing (mNGS) and pathogen-targeted NGS (ptNGS). In this study, we aimed to evaluate the performance of ptNGS in pathogen detection in CSF.

Methods: CSF specimens were acquired from 38 patients with meningitis who were diagnosed at Xuanwu Hospital, Capital Medical University between October 2020 and February 2021. DNA was extracted from the CSF samples, and pathogens were identified using both ptNGS and mNGS. SPSS 22.0 software was used to compare the pathogen detection performance of ptNGS and mNGS in CSF.

Results: Among the 38 patients with meningitis, 14 had a non-infectious disease (NID) and 24 had an infectious disease (ID). Of the 38 samples, both ptNGS and mNGS detected 9 (23.7%) positive samples, and 12 (31.6%) negative samples. Thirteen (34.2%) samples were detected to be positive by ptNGS only, and 4 (10.5%) were detected to be positive by mNGS only. The positivity rate detected by ptNGS for the ID group was higher than that detected by mNGS (P=0.080), and the positivity rates detected by ptNGS and mNGS for the NID group were comparable. The positive predictive value (PPV) and negative predictive value (NPV) of diagnosing an ID by ptNGS were 77.3% and 56.3%, respectively. While, the PPV and NPV of diagnosing an ID by mNGS were 76.9% and 44.0%, respectively. ptNGS increased the sensitivity rate by approximately 70%. The sensitivity rate of ptNGS was higher than that of mNGS (70.8% vs. 41.7%), while the specificity rate of mNGS was higher than that of ptNGS (78.6% vs. 64.3%). Additionally, ptNGS required a shorter time for pathogen diagnosis (15 vs. 24 hrs) and had lower costs than mNGS.

Conclusions: ptNGS has a number of advantages over mNGS, including its sensitivity, timeliness, and economy, all factors that are important considerations in clinical use.

 $\textbf{Keywords:} \ \ Meningitis; cerebrospinal \ fluid \ (CSF); pathogen; next-generation \ sequencing \ (NGS)$

Submitted Sep 17, 2021. Accepted for publication Nov 22, 2021.

doi: 10.21037/atm-21-5488

View this article at: https://dx.doi.org/10.21037/atm-21-5488

 $^{^{\}wedge}\ ORCID:\ Daiquan\ Gao,\ 0000-0002-2091-7567;\ Yongqiang\ Hu,\ 0000-0001-6763-0072;\ Xuebin\ Jiang,\ 0000-0002-8412-155X;\ Hao\ Pu,\ 0000-0001-9404-5549;\ Zhendong\ Guo,\ 0000-0001-6763-0072;\ Yunzhou\ Zhang,\ 0000-0001-9056-9505.$

Introduction

Meningitis is an inflammation of the fluid and meninges surrounding the brain and spinal cord. Meningitis, especially acute meningitis, represents a threat to people's health. Clinically, the cause of acute meningitis cases is not identified in approximately 50% of patients (1-3). A failure to make a timely diagnosis in patients with meningitis contributes to poor patient outcomes. Accurate information about important etiological agents is necessary to ensure appropriate management. Viral infections, followed by bacterial infections, are the most common causes of meningitis. Conversely, fungal and parasitic infections rarely cause meningitis. Non-infectious causes, such as chemical reactions, drug allergies, cancer, and inflammatory diseases (e.g., sarcoidosis) can also cause meningitis.

The cerebrospinal fluid (CSF) culture has been widely used to diagnose meningitis (4). However, for cases with negative CSF cultures, the diagnosis of bacterial meningitis may be established by positive blood cultures (5). Due to the low sensitivity of CSF cultures, several alternative methods have been introduced for pathogen detection, including polymerase chain reaction (PCR), film array, immunoassay, microarray, and next-generation sequencing (NGS) (6-9). Among these methods, immunoassay shows best timeliness; PCR and film array show best accuracy. But the throughput of all these methods are low. NGS base methods show much higher throughput. It can also detects not only pathogens but also their drug-resistant genes in on run. Metagenomic NGS (mNGS) is an increasingly rapid and comparatively low-cost means of screening CSF in an unbiased manner for a broad range of human pathogens (10,11). mNGS is a single diagnostic test with promising potential for clinical application (10,11). Recently, the new method of pathogen-targeted NGS (ptNGS) was developed and has been used to identify pathogens in respiratory tract infection or mycobacterium infection cases (12,13). At present, it seems that ptNGS has the advantages of detection sensitivity not affected by human genome and background bacteria, detection sensitivity not affected by pathogen genome size, lower detection cost, reduced sample transportation requirements, and quantitative detection of pathogens. However, its pathogen detection performance in CSF remains unclear. In this study, ptNGS and mNGS were compared with conventional test results respectively, and then ptNGS and mNGS technologies based on the NGS platform were compared horizontally. This project is the first to apply ptNGS to the pathogenic diagnosis of CSF

in patients with central nervous system infection.

In the present study, we aimed to evaluate the performance of ptNGS in pathogen detection in CSF. CSF samples from 38 patients with meningitis, who had been diagnosed using the traditional diagnostic method of CSF, were examined using the mNGS and ptNGS methods. The performance of the ptNGS and mNGS was then compared. We present the following article in accordance with the MDAR reporting checklist (available at https://dx.doi.org/10.21037/atm-21-5488).

Methods

Patients and samples

CSF specimens were acquired from 38 patients who were diagnosed with meningitis using the traditional diagnostic method of CSF at Xuanwu Hospital, Capital Medical University between October 2020 and February 2021. Based on their final diagnoses, the patients were categorized into an infectious disease (ID) group and a non-infectious disease (NID) group. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Ethics Committee of Xuanwu Hospital, Capital Medical University [No. (2020)103]. All participants provided written informed consent. The CSF specimens were stored at –20 to –80 °C awaiting use.

DNA extraction

The CSF samples were centrifuged at 13,000 rpm for 10 min and were ground on a Grinding Mill (Tiss-24, Jingxin, Shanghai, China) at 60 HZ for 10 min. The ground samples were then used for DNA/RNA extraction and purification (ZymoBIOMICS DNA Miniprep Kit, R2002) in accordance with the manufacturer's instructions. The extracted DNA was used to construct a library. Reduce host DNA contamination from clinical CSF samples: (I) before nucleic acid extraction: osmotic dissolution lyse human cells to retain the intact pathogen; the released human DNA is treated and degraded with DNase or chemical reagents (such as monoaziridin), and then the microbial nucleic acid is extracted (QIAamp DNA Microbiome Kit); (II) after nucleic acid extraction: after the nucleic acid of the specimen is extracted, the anti-methylated DNA specific binding protein antibody is used to selectively bind and remove the methylated DNA of CpG island (NEBNext Microbiome DNA Enrichment Kit).

ptNGS

A total of 544 meningitis associated pathogens were included in the identification model for meningitis diagnosis. First, the DNA sequences used in the identification model were selected as the targeted fragments. The primers were designed and synthesized at Pathogeno Biotech and Sangon Biotech (Shanghai, China), respectively. The targeted gene sequences were amplified and enriched by multiplex PCR (as described previously) (11). A sequencing library was constructed by adding the sequencing connectors to the purified PCR products and using DNA purification magnetic beads for purification. Targeted gene sequencing (300 cycles) was performed on a MiSeq system (Illumina, Inc., San Diego, CA, USA) using the MiSeq reagent kit v2. FastQ files were generated by the MiSeq Reporter software. Offline data generated by the MiSeg system were identified and counted by an adapter, and reads with a double-end length >60 bp were retained. Among the high-quality data, reads with <60 bp at either end, single-end primer recognition, or non-specific primer binding, were re-examined and deleted. Clean read pairs were obtained for identification and sequence alignment. Before aligning to pathogen reference sequences, clean read pairs were first aligned to human reference genome. Read pairs aligned to human genome were dropped to reduce host DNA contamination. Remained read pairs were aligned to pathogen reference sequences. Read paired counts for each pathogen were generated for further analysis.

ptNGS quality control: (I) nucleic acid purity control: The purity of the extracted nucleic acid is measured by Nanodrop. Generally, A260/A280 is greater than 1.8, and A260/A230 greater than 2.0 is considered a pure sample; (II) error-proof label control: every A unique error-proof label is added to each sample at the beginning of nucleic acid extraction. After the sequencing is completed, the correct and stable error-proof label signal in each sample is analyzed to monitor sample confusion or extraction/ amplification failure; (II) library quality control: the normal library fragment size is about 350 bp, there is no obvious dimer or non-specific band, and the library concentration is greater than 1 ng/µL; (IV) sequencing quality control: after filtering and comparing and analyzing the total data volume of a single library, the effective data volume should not be less than 0.08 M reads.

mNGS

The mNGS method has been described previously (8). Briefly, DNA libraries were constructed using an end-repair method; adapters were added overnight, and amplified by PCR before the analysis using an Ion Torrent Proton Sequencer (Life Technologies, Carlsbad, CA, USA). Before sequencing, the quality of the DNA libraries was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) in conjunction with a quantitative PCR. The qualified DNA libraries were prepared by emulsion PCR in the OneTouch system and sequenced on the Ion Torrent Proton (Life Technologies, South San Francisco, CA, USA) sequencing platform. High-quality sequencing data were obtained by removing low-quality and short reads (length <35 bp), and Burrows-Wheeler alignment was then used to subtract human host sequences. After removing low-complexity reads, the remaining data were then classified by aligning them to 4 microbial genome databases (including viruses, bacteria, fungi and parasites) simultaneously. RefSeq contains 4,189 wholegenome sequences of viral taxa, 2,328 bacterial genomes or scaffolds, 199 fungi associated with human infection, and 135 parasites related to human diseases.

Statistical analysis

SPSS 22.0 software was used for the data analysis. A Pearson chi-squared (χ^2) test or a Fisher's exact test was used to compare the frequency of the categorical data. A *t*-test (normal distribution) or a non-parametric test (Mann-Whitney test, non-normal distribution) was used to compare the quantitative data between the two groups. For all presented data, a P value <0.05 was considered statistically significant.

Results

Patient characteristics

The clinicopathologic information of 38 patients with meningitis is listed in Table S1. Notably, 14 patients had a NID, and 24 had an ID. Specifically, there was 18 cases of viral meningoencephalitis (VM), 3 cases of purulent meningitis (PM), 2 cases of tuberculous meningitis (TBM), and 1 case of fungal meningitis (FM) patients. Six (42.9%) and 12 (50.0%) male patients had a NID and an ID, respectively. The median age of patients with a NID and an ID was 39 (range, 11–77) and 46 (range, 15–68) years old,

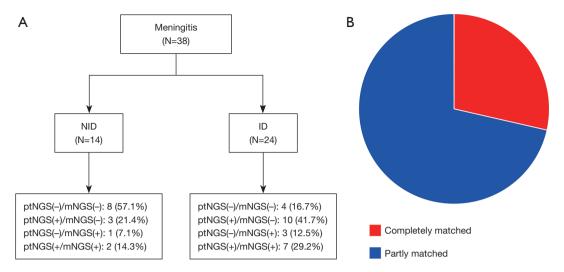


Figure 1 Overview of sample classification and comparison. (A) Based on diagnoses obtained using the traditional diagnostic method of CSF, the samples were categorized into ID or NID groups. All samples were examined by ptNGS and mNGS for the concordance analysis, and ID and NID patients were used to evaluate their diagnostic performance. (B) For the double-positive subset in the ID group (n=7), 2 patients had completely matched results and 5 had partially matched results. CSF, cerebrospinal fluid; ID, infectious disease; NID, non-infectious disease; ptNGS, pathogen-targeted next-generation sequencing; mNGS, metagenomic next-generation sequencing.

Table 1 The consistency between ptNGS and mNGS

	ptNGS (+)	ptNGS (-)	Total	Kappa	P value
mNGS (+)	9	4	13	0.148	0.490
mNGS (-)	13	12	25		
Total	22	16	38		

The symbol "+" stands for "positive", and the symbol "-" stands for "negative". ptNGS, pathogen-targeted next-generation sequencing; mNGS, metagenomic next-generation sequencing.

respectively.

Consistency between ptNGS and mNGS detection

Both ptNGS and mNGS were used to detect the pathogens in each patient. The pathogen results obtained by ptNGS and mNGS are presented in Table S1. As Figure 1A shows, of the 38 cases, both ptNGS and mNGS identified 9 (2 NID and 7 ID; 23.7%) positive samples and 12 (8 NID and 4 ID; 31.6%) negative samples. ptNGS identified 13 (34.2%) positive samples only (3 NID and 10 ID), and mNGS identified 4 (10.5%) positive samples only (1 NID and 3 ID). The consistency between ptNGS and mNGS among all the enrolled patients was estimated, and the Kappa value was 0.148 (P=0.490; see Table 1). In the ID

group, ptNGS and mNGS both identified positive samples in 7 patients (2 results were completely matched and 5 were partially matched) (see Table S2 and *Figure 1B*).

Comparison of the diagnostic performance of ptNGS and mNGS

In the ID group (n=24), 17 (70.8%) patients were diagnosed with confirmed pathogens by ptNGS, while only 10 (41.7%) were diagnosed by mNGS (see *Figure 2A*).

In the ID group, the positivity rate detected by ptNGS for was higher than that detected by mNGS (P=0.080). In the NID group (n=14), 5 (35.7%) patients were diagnosed with confirmed pathogens by ptNGS and 3 (21.4%) by mNGS. The positivity rates for the NID group between ptNGS and mNGS were comparable (see *Figure 2B*).

The performance of ptNGS and mNGS in pathogen detection was compared (see *Table 2*). The positive predictive value (PPV) and the negative predictive value (NPV) of diagnosing ID by ptNGS were 77.3% and 56.3%, respectively. Conversely, the PPV and the NPV of diagnosing ID by mNGS were 76.9% and 44.0%, respectively. ptNGS increased the sensitivity rate by approximately 70%. The sensitivity rate of ptNGS was higher than that of mNGS (70.8% *vs.* 41.7%), while the specificity rate of mNGS was higher than that of ptNGS

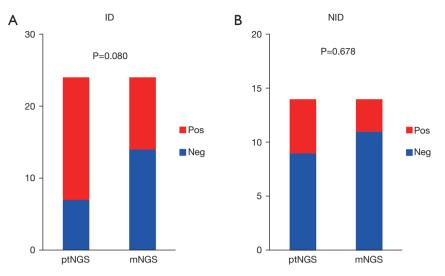


Figure 2 Positivity rate comparison between ptNGS and mNGS. (A) The differences between ptNGS and mNGS in infectious disease (n=24). (B) The differences between ptNGS and mNGS in non-infectious disease (n=14). ptNGS, pathogen-targeted next-generation sequencing; mNGS, metagenomic next-generation sequencing; ID, infectious disease; NID, non-infectious disease.

Table 2 The pathogen detection performance of ptNGS and mNGS $\,$

Method	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
ptNGS	70.8	64.3	77.3	56.3
mNGS	41.7	78.6	76.9	44.0

ptNGS, pathogen-targeted next-generation sequencing; mNGS, metagenomic next-generation sequencing; PPV, positive predictive value; NPV, negative predictive value.

(78.6% vs. 64.3%).

It took approximately 15 and 24 hrs for pathogen diagnosis by ptNGS and mNGS, respectively. mNGS has the ability to detect unknown pathogens; however, ptNGS only detects major pathogens with a clinical incidence of more than 95%. ptNGS can quantify pathogens at the copy number level, while mNGS can only perform relative quantifications. ptNGS can detect 50 copies of pathogens per milliliter of CSF sample, while the detection limit of mNGS depends on the context of the samples. *Table 3* provides a summary of the comparison of the technical advantages of ptNGS and mNGS.

Discussion

Meningitis may be infectious or non-infectious. The main types of infectious meningitis include VM, PM, TBM, and FM. Prognosis and treatment vary for different types of meningitis. Bacterial meningitis is a serious ID that can be fatal to both children and adults. The incidence and mortality rates of bacterial meningitis vary depending on the type of pathogen (14). To effectively treat bacterial meningitis, the microorganisms and their antibiotic susceptibility patterns need to be rapidly identified.

Currently, the CSF culture is the gold standard for diagnosing bacterial meningitis. However, the low bacterial growth rates cause high false negative results (15-17). Thus, new test methods need to be developed. mNGS, which is a rapid and high-throughput method for pathogen detection, has been applied to CSF samples in several studies (6-8). Miao et al. systematically compared mNGS and the CSF culture, and found that mNGS was advantageous in several aspects (10). mNGS has a higher sensitivity for pathogen identification and is less affected by previous exposure to antibiotics; thus, it has become a promising technology for the detection of IDs. However, it is difficult for mNGS to define specific microbial profiles that are diagnostic or predictive of disease development. Additionally, mNGS typically requires at least 20 million reads for each sample library and has high costs. Therefore, a new method is needed to solve these limitations. ptNGS requires only 0.08 million reads for a single sample library, which greatly improves the detection throughput and reduces the sequencing costs. Based on the superior performance of mNGS in pathogen identification, we evaluated the performance of ptNGS by comparing it to that of the mNGS.

Table 3 Comparison of the technical advantages of ptNGS and mNGS

Items	ptNGS	mNGS
Detection target	Targeted pathogens	Metagenome
Detection period	15 hrs	24 hrs
Detection of unknown pathogens	No	Yes
Limit of detection	50 copies per mL	Context dependent
Quantitative detection	Copy number level quantification	Relative quantification
Economic cost	Low	High

ptNGS, pathogen-targeted next-generation sequencing; mNGS, metagenomic next-generation sequencing.

Li et al. reported that compared with conventional tests, ptNGS (targeted amplicon sequencing assay) has better sensitivity and specificity for the diagnosis of alveolar lavage fluid bacterial and viral pathogens in severe communityacquired pneumonia patients (17). Chao et al. revealed that using pathogen target sequencing method to detect pathogenic bacteria in the sputum of patients with Acute lower respiratory infections is more accurate and sensitive than traditional Sputum culture detection (14). In this study, we use CSF samples to compare the differences between ptNGS and mNGS in the detection of meningitis patients. The analysis showed that ptNGS had a higher sensitivity rate than that of the mNGS (70.8% vs. 41.7%). However, the specificity of ptNGS was low (56.3%). This may be due to the false negative results of traditional diagnostic methods. Surprisingly, the positivity rates of ptNGS and mNGS were not higher than traditional diagnostic methods (68.4%/55.3% vs. 63.2%). This may be due to the relatively small size of the study cohort. Additionally, compared to mNGS, ptNGS took a shorter time to identify pathogens (15 vs. 24 hrs) and had lower economic costs. Although this is a small sample size study and more studies need to evaluate the performance of ptNGS in clinical use, it is proven that ptNGS has a promising potential method for guiding clinical detection.

Conclusions

ptNGS has a number of advantages over mNGS, including its sensitivity, timeliness, and economy, all of which are usually considered in clinical use.

Acknowledgments

We would like to thank all the participating subjects for

their cooperation in this study.

Funding: This work was supported by the National Key Research and Development Program of China (2020YFC2005403).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://dx.doi.org/10.21037/atm-21-5488

Data Sharing Statement: Available at https://dx.doi.org/10.21037/atm-21-5488

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-5488). HP and ZG report that they are from Shanghai Pathogeno Medical Technology Co., Ltd. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Ethics Committee of Xuanwu Hospital, Capital Medical University [No. (2020)103]. All participants provided written informed consent.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International

License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

- Glaser CA, Gilliam S, Schnurr D, et al. In search of encephalitis etiologies: diagnostic challenges in the California Encephalitis Project, 1998-2000. Clin Infect Dis 2003;36:731-42.
- Glaser CA, Honarmand S, Anderson LJ, et al. Beyond viruses: clinical profiles and etiologies associated with encephalitis. Clin Infect Dis 2006;43:1565-77.
- Granerod J, Ambrose HE, Davies NW, et al. Causes of encephalitis and differences in their clinical presentations in England: a multicentre, population-based prospective study. Lancet Infect Dis 2010;10:835-44.
- Aberle SW, Puchhammer-Stöckl E. Diagnosis of herpesvirus infections of the central nervous system. J Clin Virol 2002;25 Suppl 1:S79-85.
- 5. Gray LD, Fedorko DP. Laboratory diagnosis of bacterial meningitis. Clin Microbiol Rev 1992;5:130-45.
- Fuglsang-Damgaard D, Pedersen G, Schønheyder HC. Positive blood cultures and diagnosis of bacterial meningitis in cases with negative culture of cerebrospinal fluid. Scand J Infect Dis 2008;40:229-33.
- Başpınar EÖ, Dayan S, Bekçibaşı M, et al. Comparison of culture and PCR methods in the diagnosis of bacterial meningitis. Braz J Microbiol 2017;48:232-6.
- Ben RJ, Kung S, Chang FY, et al. Rapid diagnosis of bacterial meningitis using a microarray. J Formos Med Assoc 2008;107:448-53.

Cite this article as: Gao D, Hu Y, Jiang X, Pu H, Guo Z, Zhang Y. Applying the pathogen-targeted next-generation sequencing method to pathogen identification in cerebrospinal fluid. Ann Transl Med 2021;9(22):1675. doi: 10.21037/atm-21-5488

- Díaz KMO, Piedrahíta JAA, Suárez Brochero OF, et al. Impact of the Film Array Meningitis/Encephalitis panel in adults with meningitis and encephalitis in Colombia. Epidemiol Infect 2020;148:e173.
- Miao Q, Ma Y, Wang Q, et al. Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. Clin Infect Dis 2018;67:S231-40.
- 11. Wilson MR, O'Donovan BD, Gelfand JM, et al. Chronic meningitis investigated via metagenomic next-generation sequencing. JAMA Neurol 2018;75:947-55.
- Wilson MR, Sample HA, Zorn KC, et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. N Engl J Med 2019;380:2327-40.
- 13. Li B, Xu L, Guo Q, et al. GenSeizer: a multiplex PCR-based targeted gene sequencing platform for rapid and accurate identification of major mycobacterium species. J Clin Microbiol 2021;59:e00584-20.
- 14. Chao L, Li J, Zhang Y, et al. Application of next generation sequencing-based rapid detection platform for microbiological diagnosis and drug resistance prediction in acute lower respiratory infection. Ann Transl Med 2020;8:1644.
- Trotman H, Olugbuyi O, Barton M, et al. Pneumococcal meningitis in Jamaican children. West Indian Med J 2009;58:585-8.
- Brouwer MC, Tunkel AR, van de Beek D. Epidemiology, diagnosis, and antimicrobial treatment of acute bacterial meningitis. Clin Microbiol Rev 2010;23:467-92.
- 17. Li F, Wang Y, Zhang Y, et al. Etiology of Severe Pneumonia in Children in Alveolar Lavage Fluid Using a High-Throughput Gene Targeted Amplicon Sequencing Assay. Front Pediatr 2021;9:659164.

(English Language Editor: L. Huleatt)