



HAL
open science

Performance of clinical metagenomics in France: a prospective observational study

Jacques Fourgeaud, Béatrice Regnault, Vichita Ok, Nicolas da Rocha, Émilie Sitterlé, Meryem Mekouar, Hélène Faury, Catherine Milliancourt-Seels, Florence Jagorel, Delphine Chrétien, et al.

► To cite this version:

Jacques Fourgeaud, Béatrice Regnault, Vichita Ok, Nicolas da Rocha, Émilie Sitterlé, et al.. Performance of clinical metagenomics in France: a prospective observational study. *The Lancet Microbe*, 2023, 10.1016/s2666-5247(23)00244-6 . pasteur-04320314

HAL Id: pasteur-04320314

<https://pasteur.hal.science/pasteur-04320314v1>

Submitted on 4 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Performance of clinical metagenomics in France: a prospective observational study

Jacques Fourgeaud*, Béatrice Regnault*, Vichita Ok, Nicolas Da Rocha, Émilie Sitterlé, Meryem Mekouar, Hélène Faury, Catherine Milliancourt-Seels, Florence Jagorel, Delphine Chrétien, Thomas Bigot, Éric Troadeac, Isabelle Marques, Alexandra Serris, Danielle Seilhean, Bénédicte Neven, Pierre Frange, Agnès Ferroni, Marc Lecuit, Xavier Nassif, Olivier Lortholary†, Marianne Leruez-Villet‡, Philippe Pérot‡, Marc Eloit‡, Anne Jamet‡



Summary

Background Metagenomic next-generation sequencing (mNGS) allows untargeted identification of a broad range of pathogens, including rare or novel microorganisms. Despite the recognition of mNGS as a valuable diagnostic tool for infections, the most relevant indications for this innovative strategy remain poorly defined. We aimed to assess the determinants of positivity and clinical utility of mNGS.

Methods In this observational study, we prospectively performed short-read shotgun metagenomics analysis as a second-line test (in cases of negative first-line test or when the symptoms were not fully explained by initial positive results) or as a first-line test in life-threatening situations requiring urgent non-targeted pathogen identification at the Necker-Enfants Malades Hospital (Paris, France). All sample types, clinical indications, and patient populations were included. Samples were accompanied by a mandatory form completed by the senior clinician or pathologist, on which the clinical level of suspected infection (defined as high or low) was indicated. We assessed the variables (gender, age, immune status, initial suspicion of infection, indication, and sample type) associated with mNGS pathogen detection using odds ratios (ORs) from multivariate logistic regression. Additional investigations were carried out using specific PCR or culture techniques, to confirm positive mNGS results, or when infectious suspicion was particularly high despite a negative mNGS result.

Findings Between Oct 29, 2019, and Nov 7, 2022, we analysed 742 samples collected from 523 patients. The initial suspicion of infection was either high (n=470, 63%) or low (n=272, 37%). Causative or possibly causative pathogens were detected in 117 (25%) samples from patients with high initial suspicion of infection, versus nine (3%) samples analysed to rule out infection (OR 9·1, 95% CI 4·6–20·4; p<0·0001). We showed that mNGS had higher odds of detecting a causative or possibly causative pathogenic virus on CNS biopsies than CSF samples (4·1, 1·7–10·7; p=0·0025) and in samples from immunodeficient compared with immunocompetent individuals (2·4, 1·4–4·1; p=0·0013). Concordance with conventional confirmatory tests results was 103 (97%) of 106, when mNGS detected causative or possibly causative pathogens. Altogether, among 231 samples investigated by both mNGS and subsequent specific tests, discordant results were found in 69 (30%) samples, of which 58 (84%) were mNGS positive and specific tests negative, and 11 (16%) mNGS negative and specific tests positive.

Interpretation Major determinants of pathogen detection by mNGS are immune status and initial level of suspicion of infection. These findings will contribute, along with future studies, to refining the positioning of mNGS in diagnostic and treatment decision-making algorithms.

Funding Necker-Enfants Malades Hospital and Institut Pasteur.

Copyright © 2023 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.

Introduction

Pathogen identification during emerging microorganism outbreaks and on return from international travel, including rapidly evolving viruses (eg, coronaviruses, arboviruses, influenza virus, or Ebola virus), requires broad-spectrum diagnostic tools readily available to clinicians. Additionally, immunocompromised patients can have a wide range of opportunistic pathogens (eg, herpes viruses, enteroviruses, or astroviruses) or atypical clinical presentations (eg, rubella virus granuloma).^{1–3} Although cultures and universal PCR (eg, 16S/18S rRNA

gene PCR) allow for broad-spectrum analysis of bacterial and fungal microorganisms, conventional tests in virology consist of targeted pathogen-specific nucleic acids or antigens, or antibodies detection, which limits the list of detectable viruses to the most common or foreseen pathogens. The diagnosis of parasitic diseases is also commonly based on specific molecular or serological tests, in addition to morphological analysis by microscopy.

Microorganisms detectable by direct, high-throughput, nucleic acid sequencing of clinical samples (known as

Lancet Microbe 2023

Published Online
December 1, 2023
[https://doi.org/10.1016/S2666-5247\(23\)00244-6](https://doi.org/10.1016/S2666-5247(23)00244-6)

For the French translation of the abstract see Online for appendix 1

*Contributed equally as first authors

†Contributed equally

‡Contributed equally as last authors

Université Paris Cité, FETUS, Paris, France
(J Fourgeaud PharmD, Prof P Frange MD, M Leruez-Ville MD);
Microbiology Department, AP-HP, Hôpital Necker, Paris, France (J Fourgeaud, V Ok PharmD, É Sitterlé PharmD, M Mekouar MD, H Faury MD, C Milliancourt-Seels BSc, I Marques BSc, Prof P Frange, A Ferroni PharmD, M Leruez-Ville, A Jamet MD);
Institut Pasteur, Université Paris Cité, Pathogen Discovery Laboratory, Paris, France (B Regnault MSc, N Da Rocha MSc, F Jagorel BSc, D Chrétien BSc, É Troadeac BSc, P Pérot PhD, Prof M Eloit PhD);
Bioinformatics and Biostatistics Hub, Computational Biology Department, Institut Pasteur, Paris, France (B Regnault, T Bigot PhD);
Université Paris Cité, Centre d'Infectiologie Necker-Pasteur, IHU Imagine, Hôpital Necker, Paris, France (A Serris MD, Prof M Lecuit MD, Prof O Lortholary MD);
Département de Neuropathologie Raymond Escourolle, AP-HP-Sorbonne, Groupe Hospitalier Pitié-Salpêtrière, Paris, France (Prof D Seilhean MD);
Pediatric Hematology Immunology and Rheumatology Unit, AP-HP, Hôpital Necker, Paris, France (Prof B Neven MD);
Université Paris Cité, INSERM, Institut Imagine, Laboratory of

Immunogenetics of Pediatric Autoimmune Diseases, Paris, France (Prof B Neven); Institut Pasteur, Université de Paris, INSERM U1117, Biology of Infection Unit, Paris, France (Prof M Lecuit); Institut Pasteur, National Reference Center and WHO Collaborating Center *Listeria*, Paris, France (Prof M Lecuit); Université Paris Cité, CNRS, INSERM, Institut Necker-Enfants Malades, Team Pathogenesis of Systemic Infection, Paris, France (Prof X Nassif MD, A Jamet); Institut Pasteur, Centre National de Référence Mycoses Invasives et Antifongiques, Département de Mycologie, Labex IBEID, Paris, France (Prof O Lortholary, P Pérot); Institut Pasteur, Université Paris Cité, The WOA Collaborating Center for the Detection and Identification in Humans of Emerging Animal Pathogens, Paris, France (Prof M Eloit); École Nationale Vétérinaire d'Alfort, Université Paris-Est, Maisons-Alfort, France (Prof M Eloit); Institut du Cerveau-Paris Brain Institute-ICM, INSERM U1127, CNRS UMR7225, AP-HP, Sorbonne University, Pitié-Salpêtrière Hospital, Paris, France (Prof D Seilhean)

Correspondence to: Dr Anne Jamet, Université Paris Cité, CNRS, INSERM, Institut Necker-Enfants Malades, Team Pathogenesis of Systemic Infection, F-75015 Paris, France anne.jamet@aphp.fr

See Online for appendix 2

Research in context

Evidence before this study

The aim of clinical metagenomic next-generation sequencing (mNGS) is to detect pathogens not recognised by conventional microbiological methods to improve patient diagnosis and care. Despite mNGS being recognised as a valuable diagnostic tool for infections, the best practice for clinical application for this innovative strategy remains poorly defined. In particular, the identification of the populations or clinical situations that would benefit most from this cutting-edge technology, as well as the optimal sample types for mNGS analysis, have not yet been fully identified. We searched PubMed for studies published between Jan 1, 2015, and Feb 6, 2023, using the terms “metagenomics” AND “clinical” AND “hospital” AND “diagnostic” without language restriction and filtering for reviews and case reports. We retrieved 853 articles, from which 47 reported on the prospective use of mNGS for the diagnosis of infections. The most frequently studied medical conditions were lower respiratory tract infections (n=12 studies), CNS disorders (n=11 studies), and sepsis (n=47 studies). We identified only six prospective studies (conducted in China [n=3], France [n=2], and Switzerland [n=1]) that used multiple clinical sample types and clinical conditions. Half of them involved only adults (including a previous study from our group in immunocompromised adults), whereas the other half involved only paediatric patients. The number of patients studied ranged from ten to 168. One of these studies found no viruses and another used a pipeline limiting the search to viruses only. Of the six studies, three excluded analysis of tissue samples. Although the results of these studies suggest the clinical relevance of mNGS in complementing current routine tests, the limited size and heterogeneity of populations included, and the limited number of sample types tested do not allow definitive conclusions to be drawn.

Added value of this study

To our knowledge, this is the first national prospective study of the clinical application of mNGS to be carried out without any restrictions on sample types, medical indications, or the patient population studied. The inclusion of both immunocompetent and immunocompromised patients in the same study revealed that immunocompromised status was associated with higher odds of detecting a causative pathogen by mNGS than immunocompetent status. Our results also highlight the high potential for discovery of novel viruses or medical entities in the immunocompromised patient population. By not limiting the type of sample analysed to sterile sites, and by guiding the physician to send stool, beyond the sole investigation of diarrhoea symptoms, to also investigate liver or neurological disorders, we showed that mNGS on stool was more likely to detect a causative pathogen than mNGS on sterile fluids.

Implications of all the available evidence

In combination with previous evidence, our results highlight the strength of the non-targeted mNGS approach, particularly in immunocompromised populations. We also confirm the high positivity rate of CNS tissue analysis and show the value of stool sample analysis in causal diagnosis of hepatic or neurological conditions. In this context of broad development and availability of mNGS techniques in high-income countries' microbiology laboratories, this awareness of the best target populations and samples for high mNGS performance should assist clinicians to appropriately apply this innovative test. However, further research is needed to assess the position of mNGS, whether routinely performed by the local polyvalent microbiology laboratories or subsequently by reference centres.

metagenomic next-generation sequencing [mNGS]) include viruses (detecting either viral DNA or RNA), bacteria, fungi, and parasites. Thus, mNGS on patient fluids and tissues has enabled diagnosis that could not be achieved with conventional laboratory tests.^{4,5} In case of detection of novel or unexpected microorganisms, mNGS can be combined with light or electron microscopy, organism-specific histological immunostaining, or in-situ hybridisation, or a combination of these, to provide evidence of causality. Specific probes based on the nucleotide sequences of the organism detected by mNGS would be of great value.⁶

On Oct 29, 2019, the clinical microbiology laboratory at Necker-Enfants Malades Hospital and the Institut Pasteur (both in Paris, France) established an mNGS collaborative diagnostic platform for infection diagnosis. Although the diagnostic value of mNGS compared retrospectively with positive conventional tests is well described,^{7,8} few prospective studies have covered multiple clinical sample types and clinical conditions

(appendix 2 p 7).^{2,9-13} Thus, mNGS positioning in the arsenal of diagnostic tools still warrant clarification. In this study, we aimed to evaluate, following the STROBE metagenomics guidelines,¹⁴ the determinants of positivity and clinical utility of mNGS using a large panel of samples and clinical indications in a tertiary care centre.

Methods

Study design and participants

The samples for mNGS analysis were received from Oct 29, 2019, onwards at the Necker-Enfants Malades Hospital (Paris, France). Our inclusion criteria for mNGS indication were either negativity of first-line standard microbiology techniques (SMT), complex medical condition in which the symptoms might not be fully explained by initial SMT positive results (eg, immunodepression), or life-threatening situations without specific symptoms that required urgent non-targeted pathogen identification. First-line SMT required before mNGS analysis are described in appendix 2 (p 8). In lifzzening

situations with organ biopsy collection, mNGS was used in parallel with SMT without waiting for SMT results. On the basis of the expertise gained from our findings, we suggested to clinicians that CNS biopsies be preferred over cerebrospinal fluid (CSF) when available, and recommended stool or urine analysis to investigate hepatic, neurological, dysimmune, or nephrological disorders (prolonged viral shedding observed in immunocompromised patients).

Samples were accompanied by a mandatory prescription form completed by the senior clinician or pathologist. The clinical level of suspected infection (defined as high or low) was indicated on the prescription form by the senior clinicians. Additional information requested included current antimicrobial and immunosuppressive therapies, immune status, microbiological test results, and histological findings. On the basis of this information, categorisation as high-level or low-level of suspected infection was done independently by two microbiologists (AJ and JF) from the Necker-Enfants Malades Hospital Microbiology laboratory.

We categorised the identified microorganisms into two groups: pathogens or microorganisms without known pathogenic role. Pathogens were further subdivided into causative, possibly causative, or unlikely to be responsible for the symptoms. In cases of microorganism detection, causality was established based on the consistency of symptoms with the literature. The terms causative and possibly causative are described and illustrated in appendix 2 (p 4), as are the confirmation

requirements for each pathogen (pp 9–11). Additional investigations were carried out using specific PCR or culture techniques, to confirm positive mNGS results, or when infectious suspicion was particularly high despite a negative mNGS result.

The study was a prospective observational study with no addition to standard care procedure. Patients or their legal guardians, or both, were informed by a written letter at the time of sample collection, and were given the opportunity to object to the use of the data at any time. The study was carried out in accordance with the ethical guidelines for medical research and was approved by the AP-HP centre institutional research review board on April 9, 2022 (number 00011928).

Procedures

Samples were transported to the microbiology laboratory of the Necker-Enfants Malades Hospital on dry ice and kept frozen at -80°C before testing. Nucleic extraction protocols, quality controls, and kit used for library preparation are described in appendix 2 (pp 5, 12). Library sequencing was performed on a NextSeq500 instrument (Illumina; Evry, France) with an average of 50 million reads (1×150 bp) per sample. Testing was done in batches, on a weekly basis. The bioinformatic analysis pipeline is described in appendix 2 (p 6).

We excluded samples that were redundant (as in the case of several concomitant prescriptions for the same patient) or inappropriate (as in the case of a very strong suspicion of a specific pathogen, we suggested a specific PCR analysis instead). We also excluded samples if a

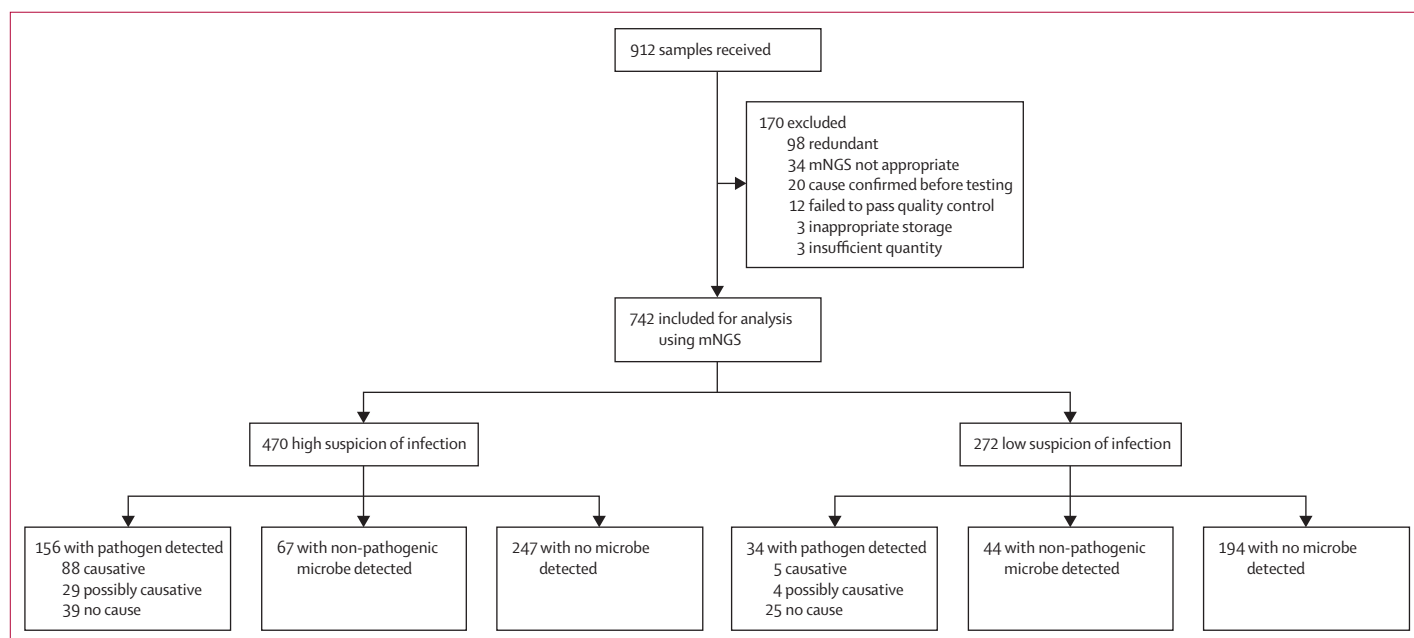


Figure 1: Flowchart of samples analysed by mNGS

The detected microorganisms were first classified according to their known pathogenicity (known pathogenic role or no known pathogenic role). The pathogenic microorganisms were then classified according to their causality in the symptoms of the studied patients. mNGS=metagenomic next-generation sequencing. SMT=standard microbiology techniques including culture of non-fastidious bacteria, broad range PCR, routine syndromic PCR panel, and specific viral PCR depending on the symptoms.

cause was identified between the time of prescription and sample processing, in cases of inappropriate storage, insufficient quantity, or quality control failure.

Outcomes

The primary outcome of this study was pathogen detection by mNGS on different clinical sample types and clinical symptoms. Secondary outcomes included effect of age, gender, immune status, and initial suspicion of infection on mNGS positivity.

Statistical analysis

Multivariate logistic regression was used to compute the odds ratio (OR) that gender, age, immune status, initial suspicion of infection, indication, and sample type has on causative pathogen detection (appendix 2 p 13); only *p* values less than 0.05 were considered significant. All analyses were conducted using GraphPad Prism software, version 9.

	Total (N=523)
Age	
Age, years	18 (9–48)
Children (≤15 years)	228 (44%)
Adults (>15 years)	295 (56%)
Gender	
Female	230 (44%)
Male	293 (56%)
Disorder types	
Neurological	204 (39%)
Hepatic	64 (12%)
Nephrological	42 (8%)
Gastrointestinal	37 (7%)
Osteomuscular	39 (7%)
Respiratory	34 (7%)
Inflammatory	34 (7%)
Cutaneous	18 (3%)
Otorhinolaryngological, ocular	17 (3%)
Haematological	17 (3%)
Fetal	9 (2%)
Cardiac	8 (2%)
Immune status	
Immunocompetent	277 (53%)
Primary immune deficiency	101 (19%)
Immunosuppressive or corticosteroid therapy	57 (11%)
Solid organ transplant	35 (7%)
Haematopoietic stem cell transplant	31 (6%)
Patients living with HIV	22 (4%)
Microorganism detection	
Causative or possibly causative pathogen	98 (19%)
Total microorganism	221 (42%)
Data are median (IQR) or <i>n</i> (%).	
Table: Demographic and baseline characteristics of patients	

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Between Oct 29, 2019, and Nov 7, 2022, we received 912 samples with a prescription to perform mNGS analysis (figure 1). We excluded 170 samples, with most exclusions (*n*=98, 58%) occurring due to redundancy—ie, when multiple samples were collected from the same patient or from the same tissue only certain samples were analysed according to the clinical situation. The excluded samples were other biological fluids or stool (*n*=57, 34%), blood (*n*=43, 25%), tissue samples (*n*=43, 25%), and CSF (*n*=27, 16%). 643 (87%) of 742 analysed samples were obtained from health facilities located in Paris (appendix 2 p 1). The median time for mNGS analysis reporting was 9 days (IQR 5–12) from sample collection to data analysis. In nine life-threatening situations, we reduced the turnaround time to 3 days. The 742 analysed samples were collected from 523 patients for whom no microbiological cause of their symptoms had been found by conventional techniques at the time of mNGS prescription (table; appendix 2 p 8). For 129 (25%) of the 523 patients, we analysed several samples from different dates or types, or both (number of samples from any of the 129 individuals ranged from 2 to 8). The median age of all 523 patients was 18 years (range 0–89, IQR 9–48). In 306 (59%) patients, infection with a pathogen was suspected and mNGS was prescribed to identify the causative pathogen or pathogens, whereas in the remaining 217 (41%) patients, the level of suspicion of infection was lower and mNGS was prescribed to rule out a possible infection (eg, before initiating immunosuppressive therapy or stopping anti-infective therapies). The number of samples received per patient was higher in the former cases (high probability of infection) than in the latter (1.5 [95% CI 1.4–1.6] vs 1.3 [1.2–1.4]; *p*=0.0401). The characteristics of the analysed samples are summarised in figure 2. All sample types were acceptable, with liver and CNS biopsies being the primary sample type analysed among tissue samples, and urine and bone marrow aspirates being the primary sample types analysed among body fluids (figure 2).

At least one microorganism (ie, viruses, bacteria, or fungi independent of their involvement in the symptoms) was detected in 301 (41%) of 742 mNGS samples, and 190 (63%) of these 301 detected microorganisms were pathogens (interpreted as causative or not). mNGS allowed causal diagnosis by identification of a causative pathogen in 98 (19%) of 523 patients.

We classified the detected microorganisms according to sample types and clinical or biological signs. The highest proportions of pathogens were detected in stool and fluids (figure 3A). 24 (46%) of 52 stool samples allowed detection

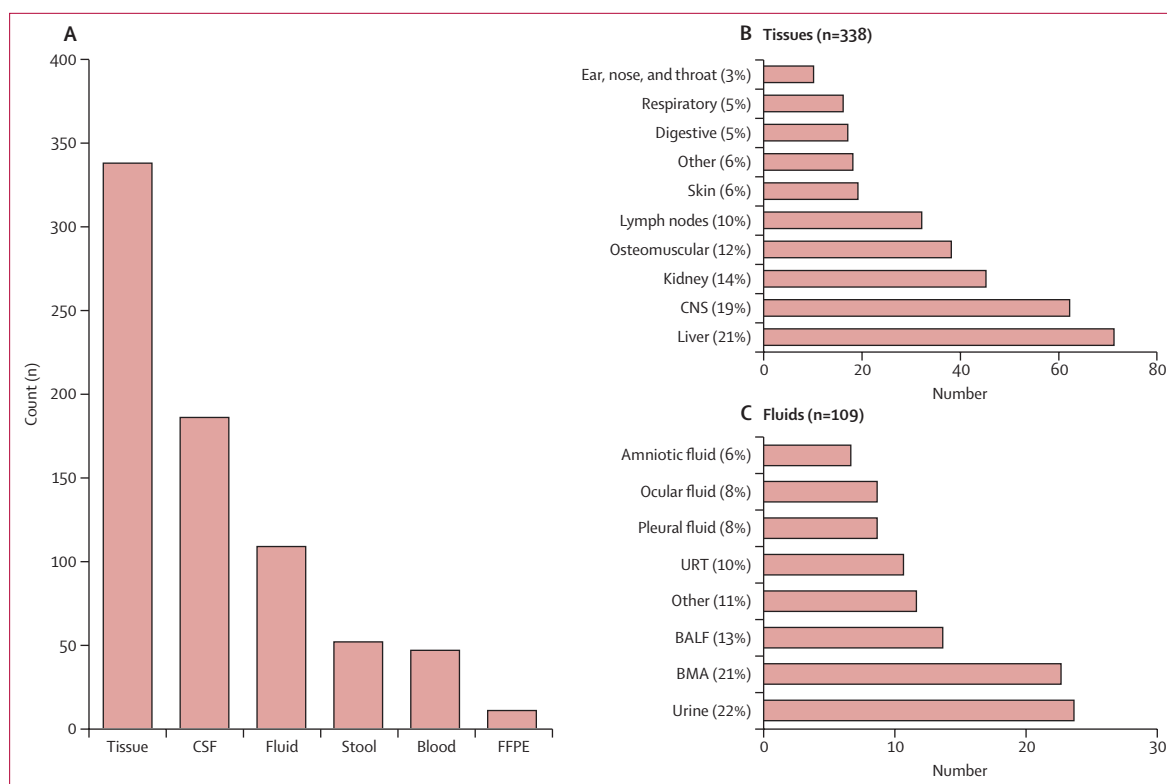


Figure 2: Characteristics of samples analysed by mNGS

(A) Distribution of sample types among the 742 samples analysed by mNGS in the present study. (B) Distribution of the sample origins in the tissue category. (C) Distribution of the sample origins in the fluid category. CSF=cerebrospinal fluid. FFPE=formalin-fixed paraffin-embedded. BMA=bone marrow aspiration. URT=upper respiratory tract. BALF=bronchoalveolar lavage fluid. mNGS=metagenomic next-generation sequencing.

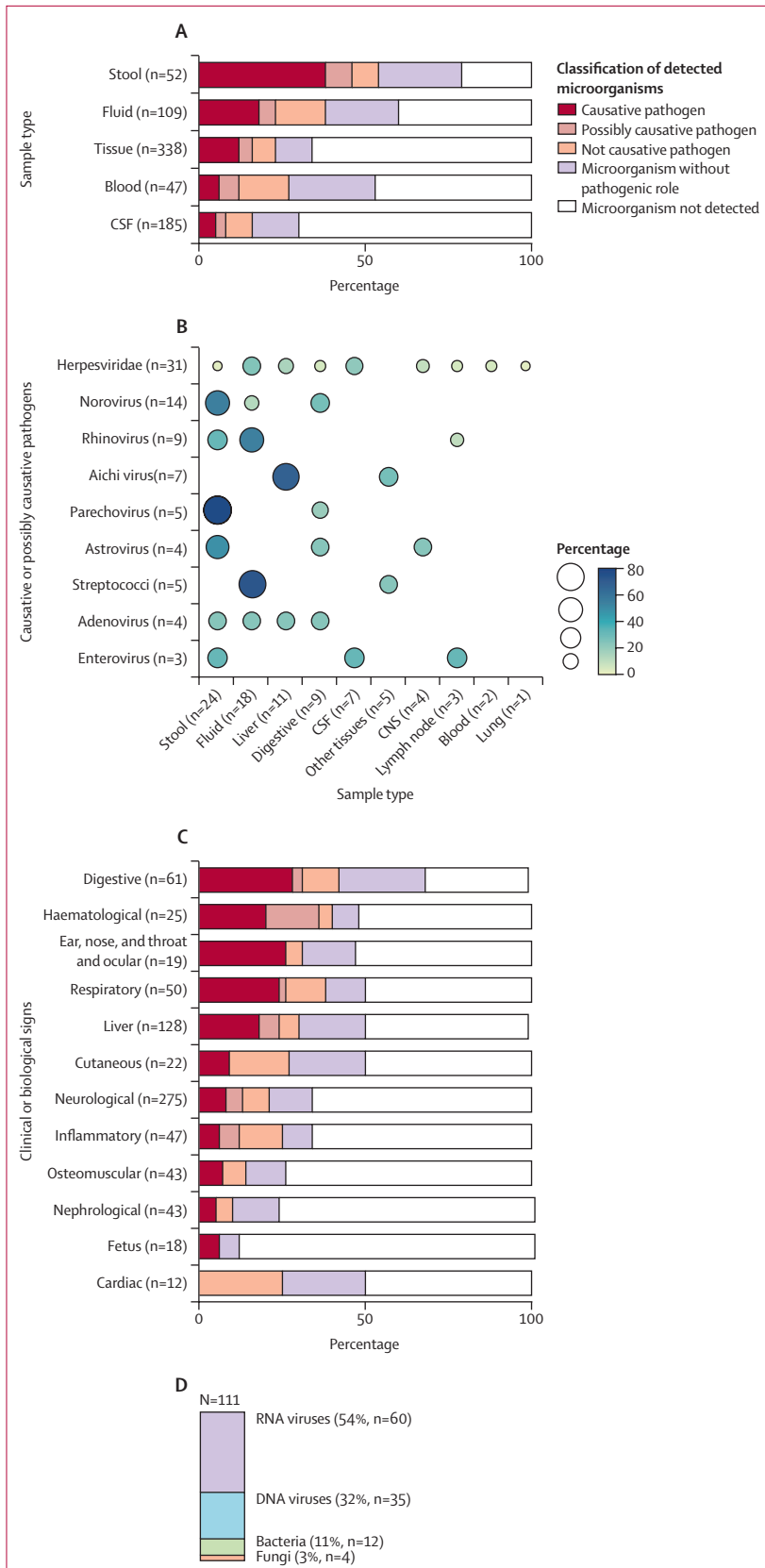
of causative or possibly causative pathogens (figure 3B; appendix 2 pp 16–17), whereas pathogens were detected in CSF in only 13 (7%) of 185 samples. The highest proportions of samples in which pathogens were identified were collected from patients with gastrointestinal, haematological, respiratory, or hepatic symptoms (figure 3C). The percentage of positivity for different sample types and clinical indications is indicated in appendix 2 (pp 14–15).

In samples in which only one pathogen was detected (n=126), we classified 111 microorganisms as causative or possibly causative pathogens, of which viruses (95 [86%]), bacteria (12 [11%]), and fungi (four [4%]) were identified (figure 3D). Viruses were mainly single-stranded RNA viruses (52 [55%] of 95) and double-stranded DNA viruses (31 [33%] of 95; figure 3). The most frequently detected microorganisms were human herpes viruses, picornaviruses (rhinoviruses, aichi viruses, parechoviruses and enteroviruses), noroviruses, astroviruses, streptococci, and adenoviruses (figure 3B). The complete list of causative or possibly causative microorganisms identified is provided in appendix 2 (pp 16–17). In 126 samples, causative or possibly causative pathogens were detected alone in 92 (73%), and associated with another pathogen (causative or not) in 34 (27%). The most common microorganisms detected in cases of co-detection of pathogens were

cytomegalovirus (n=6), noroviruses (n=4), and adenoviruses (n=4; appendix 2 pp 16–17).

To assess the variables (gender, age, immune status, initial suspected diagnostic of infection, indication of mNGS, and sample type) associated with the detection of causative or possibly causative pathogens by mNGS, we calculated ORs by multivariate logistic regression. CNS biopsies were associated with a significantly higher rate of detection of causative or possibly causative pathogens than CSF samples (14 [23%] of 63 vs 13 [7%] of 185; OR 4.1, 95% CI 1.7–11; p=0.0025; figure 4). For 15 patients, we received both CNS (n=16) and CSF (n=21) samples, of which six CNS samples and two CSF samples were positive. In two cases, both CSF and CNS samples were positive, whereas in four cases only the CNS sample was positive. Human herpes simplex virus (HSV), human herpesvirus 6 (HHV6), cytomegalovirus, *Aggregatibacter aphrophilus*, and *Enterococcus faecium* detected by mNGS in CSF were concomitantly detected in CSF by mNGS and routine standard PCR techniques. In contrast, mNGS could identify human pegivirus (HPgV) in two CSF samples in which standard PCR techniques (not targeting HPgV) were negative.¹⁵

There was no association between the category of clinical or biological indications underlying mNGS



prescriptions and an increased rate of detection of causative or possibly causative pathogens (rates ranging from zero pathogens in 12 [0%] cardiac samples to nine [36%] in 25 haematological samples; figure 4). Conversely, the initial high suspicion of infection was found to be a major determinant of the positivity rate of mNGS: causative or possibly causative pathogens were detected in 117 [25%] of 470 samples from patients with initial suspected diagnostic of infection, versus nine [3%] of 272 samples analysed to rule out the presence of a pathogen (OR 9.1, 95% CI 4.6–20.4; $p < 0.0001$). Causative or possibly causative pathogens were identified more frequently in samples from immunodeficient patients compared with immunocompetent patients (95 [25%] of 386 vs 31 [9%] of 356, respectively; OR 2.4, 95% CI 1.4–4.1; $p = 0.0013$). The type of the analysed samples and the proportion of patients with a high suspicion of infection were similar between those who were immunocompetent and those who were immunodeficient. Among immunodeficient patients, the highest rate of detection of causative or possibly causative pathogens was found in those with primary immune deficiencies (2.0, 1.1–3.8; $p = 0.0292$; figure 4).

Confirmatory tests include existing conventional laboratory techniques that were not performed in the first instance due to lack of initial suspicion (eg, detection of Epstein-Barr virus, cytomegalovirus, enteroviruses) or dedicated tests developed based on mNGS findings or literature (Aichi virus, HPgV, Astrovirus MLB1, or VA1). It was possible to perform specific confirmatory tests in 106 (84%) of 126 cases in which causative or possibly causative pathogens were identified by mNGS. Concordance with specific confirmatory tests was 100% (60 of 60) for RNA viruses, 94% (29 of 31) for DNA viruses, and 91% (ten of 11) for bacteria. The confirmation requirements for each pathogen are detailed in appendix 2 (pp 9–11).

In 125 samples, the possibility of infection was deemed particularly high and the mNGS analysis did not identify any causative or possibly causative pathogens; therefore, additional investigations were conducted using specific PCR techniques (searching for Epstein-Barr virus, cytomegalovirus, HSV-1, HHV-6, hepatitis B virus, polyomaviruses, HPgV, rhinovirus, HIV-1, SARS-CoV-2, enteroviruses, Aichi virus, *Tropheryma whipplei*, *Mycobacterium* spp 16S, and *Candida* spp 23S), serology (for cytomegalovirus or HSV-1), or prolonged culture (for mycobacteria). Altogether, among 231 samples investigated

Figure 3: Classification of detected microorganisms

(A) Classification of detected microorganisms according to sample types. (B) The most prevalent families of causative or possibly causative pathogens identified, and their distribution among the different sample types. (C) Classification of detected microorganisms according to clinical or biological signs leading to mNGS prescription. (D) Classification of causative or possibly causative pathogens detected in mono-infection. The total for each row sums to 100%. mNGS=metagenomic next-generation sequencing. CSF=cerebrospinal fluid.

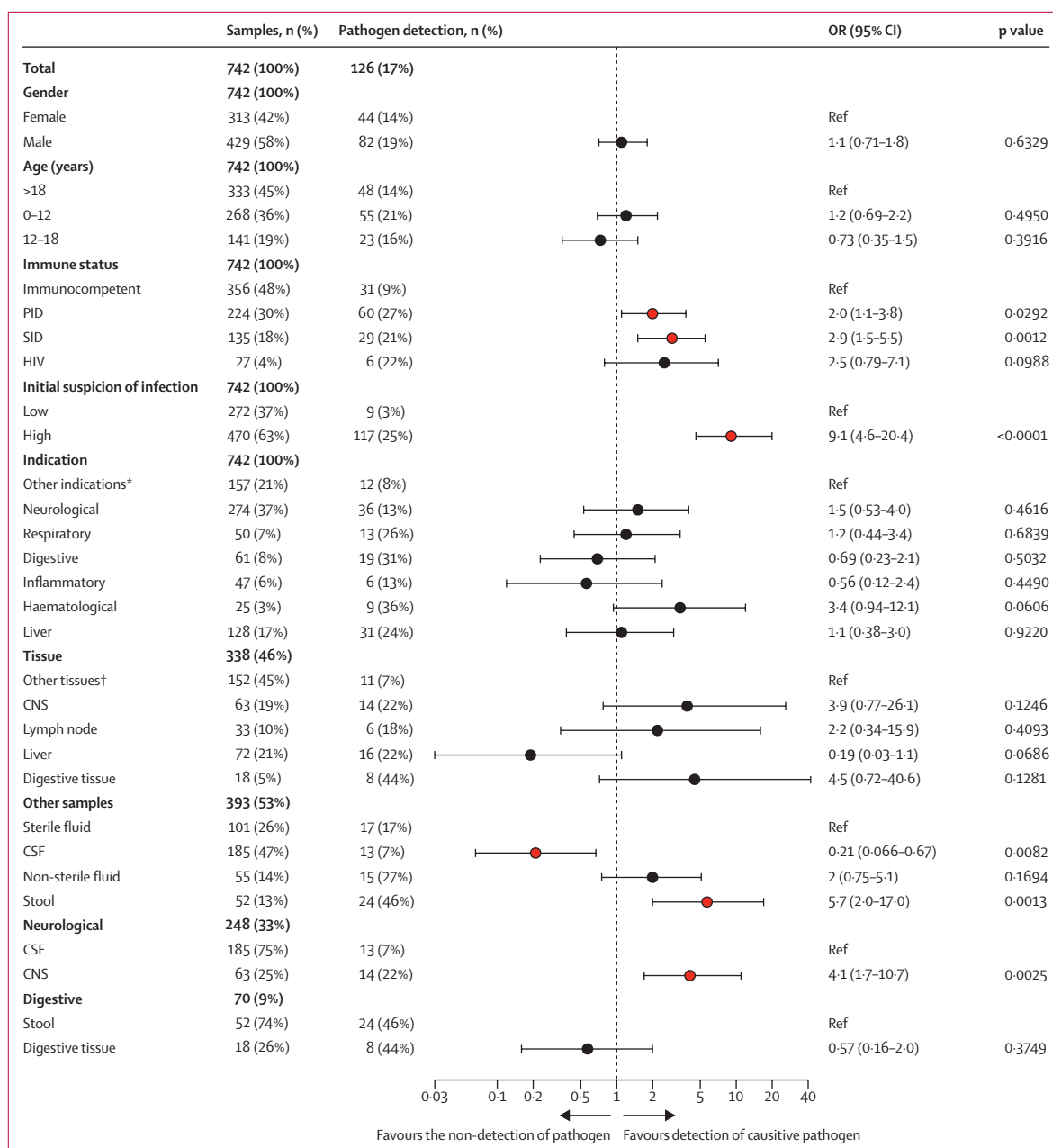


Figure 4: Causative or possibly causative pathogen detection

Association among categorical variables (gender, age, initial suspicion of infection, indication, immune status, and sample type) and causative pathogen detection was assessed by multivariate logistic-regression analysis. Only p values less than 0.0500 were considered significant and are indicated by red dots. PID=primary immune deficiency, including agammaglobulinemia, severe combined immunodeficiency, combined immunodeficiency. SID=secondary immune deficiency, including dysimmune condition post-haematopoietic stem cell transplantation, or immunosuppressive therapy. HIV=patients living with HIV. CSF=cerebrospinal fluid. OR=odds ratio. *Nephrological, osteomuscular, cardiac, fetal, dermatological, otorhinolaryngological, and ocular disorders. †Lung, kidney, osteo muscular, skin, ear, nose, throat, eye, spleen, and placenta.

by both mNGS and subsequent specific tests, discordant results were found in 69 (30%) samples, of which 58 (84%) were mNGS positive and specific tests negative, and 11 (16%) mNGS negative and specific tests positive. The 11 cases with lack of findings with mNGS and positive results of conventional tests are described in appendix 2

(p 18). Among the 58 samples with results considered as false positive (specific tests negative), the number of reads detected with mNGS for the suspected pathogen was below five in 56 (97%) of 58 samples, including contamination in the same sequencing run by a high viral load of Norovirus GII (n=3) and unconfirmed SARS-CoV-2 sequences (n=11).

The two remaining samples included 23 reads of Epstein-Barr virus detected in CSF, and 15 reads of BK virus detected in a liver biopsy sample.

mNGS performed in amniotic fluid of a pregnant woman referred for fetal brain abnormalities allowed us to detect three Zika virus (ZIKV)-specific sequences (two from the *envelope* gene and one from the *NS1* region) of the Asian genetic lineage sharing 96.1% nucleotide sequence identity with ZIKV French Polynesia/HPF/2013/KJ776791, which is consistent with the patient's travel in Thailand.¹⁶ We also detected 40 sequences of Dengue virus type 1 in the urine sample of a child (aged 7 years) from French Guiana who had received a haematopoietic stem cell transplant and presented with neurological symptoms. Vaccine rubella strain was identified in a skin granuloma in a patient (aged 20 years) with primary immunodeficiency, a previously reported, but rare, disease.³ We also typed 29 gastroenteritis viruses (human adenovirus 1, Aichi virus genotype B, human astrovirus 1, astrovirus MLB1 or VA1, norovirus GI.6 or GII.4, and sapovirus G.I or G.V), and 14 respiratory viruses (human respiratory syncytial virus A; rhinovirus A, B, or C; and Enterovirus D68 or Coxsackievirus A4 or B5). Finally, our pipeline allowed us to identify a lethal case of encephalitis with European bat lyssavirus 1 in cerebral tissue⁴ and a novel human hepatitis virus (a circovirus) in a liver sample (HCirV-1).⁶

Regarding bacterial pathogens, we were able to genotype the *S aureus* strain present in respiratory fluid from a patient with cystic fibrosis who died from acute pneumonia and showed that the strain belonged to complex clonal 30 and harboured virulence-associated superantigenic toxin genes (*sea* and *tst-1*).¹⁷ In a case of chronic diarrhoea in an immunocompromised patient, we were able to infer the serotype of the *Salmonella* strain present in the patient's stool (ie, serotype *Typhimurium*).

Discussion

In this study, we report the prospective use of mNGS during the first 3 years of its implementation in our centre. The addition of mNGS to our diagnostic workflow resulted in a causal diagnosis in 117 (25%) of 470 samples when the main diagnostic hypothesis at the time of sampling was infection, and did not detect any pathogen in 263 (97%) of 272 samples for which mNGS was prescribed to rule out infection. Overall, results of this study show the potential benefit of adding mNGS to routine diagnosis workflows for detection and discovery of rare or novel pathogens.^{4,6,16,18} We identified major determinants of clinical benefits, especially in immunocompromised patients and in individuals with available brain biopsies or stool samples.

Our study primarily identified viruses (95 [86%] of 111 infections diagnosed by mNGS in our cohort, including 57 [60%] of 95 RNA viruses), consistent with the second-line diagnostic approach where most causative bacteria, parasites, and fungi have been ruled out through

microscopic examination, before culture and broad-range PCR (16S, 16–23S, and 18S). Several drugs with broad-spectrum antiviral activities—some of which are still in development, such as ribavirin, favipiravir, cidofovir, and brincidofovir—might be helpful to treat severe viral infections for whom no specific treatment has been validated to date.¹⁹ Finally, the management of immunocompromised patients and the identification with mNGS of new clinical entities in this particularly infection-prone population are of great value.^{18,20} In an immunocompromised patient population, we observed three situations in which metagenomics had a benefit as an untargeted and sensitive diagnostic tool: (1) in case of opportunistic infections restricted to immunodeficient patients (eg, encephalitis due to astrovirus VA1 in patients with X-linked agammaglobulinemia);¹ (2) in case of infection with microorganisms known to be pathogenic in immunocompetent patients, but with atypical symptoms in terms of chronicity, severity, or clinical presentation (eg, skin granuloma triggered by rubella virus), or a combination of these;²¹ and (3) in cases where infectious symptoms might be confused with inflammatory or dysimmune manifestations (eg, enteric viruses causing symptoms versus gastrointestinal graft versus host disease).²² Overall, in 263 cases, the lack of pathogen detection by mNGS contributed to rule out infection before clinicians started immunosuppressive therapy.

In our setting, we observed that false-positive mNGS results for pathogen detection were associated with a very low number of reads (fewer than five), prompting us to confirm the mNGS results with specific techniques before reporting results to the clinician. False-positive mNGS results probably included contamination in the same series by high load of microorganism in another sample or contamination from other series, as observed for SARS-CoV-2 sequence detection during the COVID-19 outbreak. Alternatively, it could be true positive because of a greater sensitivity of mNGS, as shown in a case of congenital Zika virus infection (appendix 2 p 2).¹⁶ False negatives correspond to low load of pathogens. Loads of viruses detected only by qPCR were very low or non-pathogenic reactivation of herpes viruses. Two microorganisms were detected only in culture (not detected by PCR), in line with the detection limit by molecular approach (appendix 2 p 18).

Most studies have evaluated the diagnostic performance of mNGS by comparing it with conventional microbiological tests based on retrospective sample analysis.²³ Those studies suggest positioning mNGS as a first-line test would increase the diagnosis rate compared with SMT. Owing to the current cost of mNGS compared with standard microbiology tests, we have initially decided to use it mostly as a second-line microbiological test. In our centre, clinical microbiologists help in the prescription of the mNGS and can thus postpone it by initially performing other conventional tests or, conversely, proposing it as first line in case of tissue

samples. In 65 (9%) of 742 cases, the prescriptions came from pathologists who suspected an infectious process based on histological findings (eg, lymphocytic infiltrate or placental intervillitis). Samples addressed by pathologists, allowed us to diagnose a case of European bat lyssavirus 1 encephalitis in western Europe,⁴ and a series of unexplained multiorgan involvement due to Aichi virus.²⁰

Our study has some limitations. We focused on the determinants of mNGS positivity and we did not assess the effect of mNGS on the management of patients, nor did we include a cost analysis. We revealed the potential contribution of mNGS stool analysis to the diagnosis of hepatic,¹⁸ neurological disorders, or unexplained inflammatory diseases in immunocompromised hosts. Since pathogen excretion in the stool can be prolonged, and many individuals have low levels of pathogens in the stool, it is difficult to show causality beyond digestive symptoms. Further studies, including patient follow-up with monitoring of pathogen load kinetics and organ biopsies are, therefore, warranted to establish guidelines on the basis of how mNGS fits into diagnostic and treatment decision making algorithms. Another limitation is the calculation of turnaround time from nucleic acid extraction. The distribution of time intervals between sampling and mNGS prescription, which will be highly variable (ranging from immediate to several weeks when prescribed by a surgeon or a pathologist, respectively), would be interesting information. Finally, a delay of 3–9 days to obtain mNGS results might seem long compared with rapid, targeted PCR tests available in a matter of hours. However, non-targeted tests such as 16S PCR or uncommon specific PCRs or serologies performed by reference centres can take at least several days.

In the coming years, mNGS will probably allow the discovery of underdiagnosed or emerging pathogens. The increasing number of immunocompromised patients will most probably correlate with the emergence of new pathogen variants. Two scenarios might then occur. First, the mNGS could remain a second-line test for referral laboratories, while the most common newly discovered agents are progressively added to the first-line detection panels. For instance, following our mNGS findings, we have implemented a specific PCR for routine detection of Aichi virus, astrovirus VA1, and circovirus HCirV-1 in our centre and currently perform it as a first-line test before mNGS in evocative cases. The development of very broad range first-line tests targeting hundreds of pathogens, by combining specific amplification or capture systems and NGS, could be useful in this scenario. Second, the democratisation of mNGS, through its implementation in an increasing number of laboratories using high-throughput sequencing, could allow its use as a first-line test progressively supplanting the targeted approaches in all laboratories. In conclusion, the results of this prospective

observational study will contribute, along with future studies, to refining the positioning of mNGS in diagnostic and treatment decision making algorithms.

Contributors

AJ, JF, PP, BR, and ME conceptualised the study. AJ, JF, BR, TB, PP, and ME curated the data. JF, BR, NDR, FJ, DC, TB, ÉT, PP, ME, and AJ did the formal analysis. All authors contributed to the methods. AJ and JF wrote the original draft. All authors reviewed and edited the Article. AJ, JF, BR, PP, and ME have accessed and verified all the data in the study. All authors had full access to all the data of the study and accepted responsibility to submit for publication.

Declaration of interests

We declare no competing interests.

Data sharing

Requests for data sharing should be submitted to AJ and PP for consideration. Access to anonymised data might be granted following review via the AP-HP centre institutional research review board.

Acknowledgments

ME is part of the STROBE-metagenomics working group. We thank the sequencing platform of Necker-Enfants Malades Hospital and Biomics platform of Institut Pasteur for providing access to the sequencers. We thank Fanny Lanternier and Claire Rouzaud for providing clinical material from immunocompromised patients. We thank the technicians, biologists, and administrative staff of the microbiology department at Necker-Enfants Malades Hospital.

References

- Brown JR, Morfopoulou S, Hubb J, et al. Astrovirus VA1/HMO-C: an increasingly recognized neurotropic pathogen in immunocompromised patients. *Clin Infect Dis* 2015; **60**: 881–88.
- Parize P, Muth E, Richaud C, et al. Untargeted next-generation sequencing-based first-line diagnosis of infection in immunocompromised adults: a multicentre, blinded, prospective study. *Clin Microbiol Infect* 2017; **23**: 574.
- Neven B, Pérot P, Bruneau J, et al. Cutaneous and visceral chronic granulomatous disease triggered by a rubella virus vaccine strain in children with primary immunodeficiencies. *Clin Infect Dis* 2017; **64**: 83–86.
- Regnault B, Evrard B, Plu I, et al. First case of lethal encephalitis in western Europe due to European bat lyssavirus type 1. *Clin Infect Dis* 2022; **74**: 461–66.
- Pérot P, Bielle F, Bigot T, et al. Identification of umbre orthobunyavirus as a novel zoonotic virus responsible for lethal encephalitis in 2 French patients with hypogammaglobulinemia. *Clin Infect Dis* 2021; **72**: 1701–08.
- Pérot P, Fourgeaud J, Rouzaud C, et al. Circovirus hepatitis infection in heart-lung transplant patient, France. *Emerg Infect Dis* 2023; **29**: 286–93.
- Gu W, Deng X, Lee M, et al. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat Med* 2021; **27**: 115–24.
- Tao Y, Yan H, Liu Y, et al. Diagnostic Performance of metagenomic next-generation sequencing in pediatric patients: a retrospective study in a large children's medical center. *Clin Chem* 2022; **68**: 1031–41.
- Chen L, Zhao Y, Wei J, et al. Metagenomic next-generation sequencing for the diagnosis of neonatal infectious diseases. *Microbiol Spectr* 2022; **10**: e0119522.
- Lamoureux C, Surgers L, Fihman V, et al. Prospective comparison between shotgun metagenomics and sanger sequencing of the 16S rRNA gene for the etiological diagnosis of infections. *Front Microbiol* 2022; **13**: 761873.
- Zhang R, Zhuang Y, Xiao ZH, et al. Diagnosis and surveillance of neonatal infections by metagenomic next-generation sequencing. *Front Microbiol* 2022; **13**: 855988.
- Shen H, Shen D, Song H, et al. Clinical assessment of the utility of metagenomic next-generation sequencing in pediatric patients of hematology department. *Int J Lab Hematol* 2021; **43**: 244–49.
- Kufner V, Plate A, Schmutz S, et al. Two years of viral metagenomics in a tertiary diagnostics unit: evaluation of the first 105 cases. *Genes* 2019; **10**: 661.

- 14 Bharucha T, Oeser C, Balloux F, et al. STROBE-metagenomics: a STROBE extension statement to guide the reporting of metagenomics studies. *Lancet Infect Dis* 2020; **20**: e251–60.
- 15 Bukowska-Oško I, Perlejewski K, Pawełczyk A, et al. Human pegivirus in patients with encephalitis of unclear etiology, Poland. *Emerg Infect Dis* 2018; **24**: 1785–94.
- 16 Fourgeaud J, Regnault B, Faury H, et al. Fetal Zika virus infection diagnosed by metagenomic next-generation sequencing of amniotic fluid. *Ultrasound Obstet Gynecol* 2023; **61**: 116–17.
- 17 Tam K, Torres VJ. *Staphylococcus aureus* secreted toxins and extracellular enzymes. *Microbiol Spectr* 2019; **7**: 72.16.
- 18 Riller Q, Fourgeaud J, Bruneau J, et al. Late-onset enteric virus infection associated with hepatitis (EVAH) in transplanted SCID patients. *J Allergy Clin Immunol* 2023; **151**: 1634–45.
- 19 Stevaert A, Groaz E, Naesens L. Nucleoside analogs for management of respiratory virus infections: mechanism of action and clinical efficacy. *Curr Opin Virol* 2022; **57**: 101279.
- 20 Fourgeaud J, Lecuit MM, Péro P, et al. Chronic aichi virus infection as a cause of long-lasting multi-organ involvement in patients with primary immune deficiencies. *Clin Infect Dis* 2023; **77**: 620–28.
- 21 Wanat KA, Perelygina L, Chen MH, et al. Association of persistent rubella virus with idiopathic skin granulomas in clinically immunocompetent adults. *JAMA Dermatol* 2022; **158**: 626–33.
- 22 Naymagon S, Naymagon L, Wong S-Y, et al. Acute graft-versus-host disease of the gut: considerations for the gastroenterologist. *Nat Rev Gastroenterol Hepatol* 2017; **14**: 711–26.
- 23 Liu J, Zhang Q, Dong YQ, Yin J, Qiu YQ. Diagnostic accuracy of metagenomic next-generation sequencing in diagnosing infectious diseases: a meta-analysis. *Sci Rep* 2022; **12**: 21032.