FIND
Diagnosis for all

Near patient pathogen sequencing

Near patient pathogen sequencing

June 2022

This report was developed by PHG Foundation for FIND. We would like to thank those who contributed to the development and review of this report:

Writers:

Chantal Babb de Villiers (PHG Foundation)

Laura Blackburn (PHG Foundation)

Hayley Carr (PHG Foundation)

Joanna Janus (PHG Foundation)

Mark Kroese (PHG Foundation)

Heather Turner (PHG Foundation)

Reviewers:

Sophie Crettaz (FIND)

Cecilia Ferreyra (FIND)

Birgitta Gleeson (FIND)

Antonio Pedrotta (FIND)

Anita Suresh (FIND)

Swapna Uplekar (FIND)

Expert consultees, and the authors and contributors to section 5.5.1 (India MicroLabs), are acknowledged in Appendix 8.4.

URLs correct as of 22 June 2022

Contents

Executive summary

Introduction

Pathogen genome sequencing is used in the management of infectious diseases, supporting diagnosis, treatment decisions, outbreak management, surveillance and monitoring of antimicrobial resistance. The majority of this sequencing takes place in centralised laboratories, with samples being transported away from the patient setting for sequencing. Recent infectious disease outbreaks, including the COVID-19 pandemic, have led to considerations around near patient approaches to laboratory services, including sequencing. Near patient pathogen sequencing (NPPS) is sequencing performed outside a centralised laboratory environment, closer to the patient, and is proposed to offer quicker services that can be adapted to local situations. Locations for NPPS can range from smaller hospital- or clinic-based laboratories to mobile and in the field laboratories.

FIND, the global alliance for diagnostics, wishes to understand the opportunities and challenges of NPPS and support its implementation, with a view to providing more flexible and timely solutions to infectious disease management, particularly in lower resource settings.

Technologies and workflows

In theory, all sequencing approaches and platforms could be used in a near patient context. Due to their portability, Oxford Nanopore Technologies' sequencers, particularly the MinION, are currently the most commonly used platform for NPPS. However, there are examples where smaller Illumina platforms – e.g. iSeq – and the ThermoFisher IonTorrent sequencer have been used in near patient or field situations.

Next generation sequencing approaches can be broadly divided into two categories based on whether they are pathogen agnostic – not aimed at sequencing any one specific pathogen – or pathogen targeted, used for sequencing predefined pathogens.

Many of the common features required of sequencing platforms to carry out NPPS in lower resource settings are similar to the requirements for sequencing in centralised yet resource limited laboratories. These are:

- Ability to operate in situations with unreliable electricity and temperature control
- Ability to operate with an unreliable cold chain
- High degree of automation in terms of workflow processes and procedures
- Access to offline analysis tools
- Low maintenance equipment with limited moving parts – this includes sequencers and accessory laboratory equipment
- Workflows that limit the amount of accessory equipment needed
- Equipment that can be moved or transported without requiring extensive set up and recalibration
- Minimal reliance on other laboratory functions, e.g. culture
- Flexibility in terms of sample number throughput

Sequencing equipment, precise method choice, and other workflow design factors including workflows specifically intended to simplify more complex approaches, will also have an impact on the overall choice of the sequencing-based assay. There is research and development in many of the areas listed above, however further and ongoing commitment is required to develop a range of solutions to support implementation of NPPS in lower resource settings.

Mobile laboratories

A mobile laboratory is a laboratory that is either fully housed within or transported by a vehicle or person to be set up and used in a non-standard (semi-permanent or temporary) laboratory structure. While not all mobile laboratories carry out near patient testing, and not all near-patient laboratories are mobile, they nevertheless provide useful exemplars for exploring how to optimise laboratory and sequencing capabilities in resource limited and near patient settings. Mobile laboratories can provide rapid and flexible response services in emergencies – for example rapid diagnostics or genomic epidemiology – or support more routine needs, such as ongoing disease surveillance.

One example is the East African Community mobile laboratory network, a regional network of nine mobile laboratories that can be transported by vehicle. A further six container laboratories with sequencing capacity, that can be transported by small lorry, are in development. A mobile laboratory example from the Ebola epidemic is a laboratory in a suitcase which is easily transported and has been used in a range of field situations, more recently MicroLabs have been established in India to carry out genomic surveillance of SARS-CoV-2.

The feasibility of mobile laboratories and the successful deployment of sequencing, including in field settings, has been demonstrated. However, there continues to be a need to improve the sequencing technology, infrastructure and protocol standardisation, and to provide training in the skills required to operate such laboratories.

Case studies

Near patient pathogen sequencing could have clinical utility across a range of diseases, as demonstrated by three case studies that highlight the current use and future potential:

Sepsis is a life-threatening condition caused by a dysregulated host immune response to infection, caused by a range of pathogens, most commonly bacteria. The priority in sepsis management is the initial recognition of possible sepsis symptoms and deciding if a patient is at high, moderate or low risk. This involves accurate and rapid diagnosis of the pathogen causing sepsis, as well as its antimicrobial susceptibility profile, to support timely treatment. This process often requires culture (for bacterial pathogens), which as a method has its limitations and is often not available in lower resource settings. Near patient sequencing is being explored as a sepsis diagnostic test – either metagenomic tests, where all genetic material in a sample is sequenced, or pathogen agnostic targeted tests, whereby pre-defined genetic targets found in a range of pathogens are sequenced. Some commercial NGS tests are available, and clinical trials are ongoing, with evidence being generated for the utility of these tests in a near patient setting.

Respiratory diseases are infections of the upper and lower respiratory tract and can be challenging to accurately diagnose due to the overlapping presentations and causes of these diseases. Accurate and timely identification of the pathogen is important to appropriately prescribe therapy, and for some lower respiratory tract infections in particular, collection of suitable samples for testing can be a challenge. Culture is considered a gold standard for diagnosis. While sequencing could improve diagnostic speed and accuracy, there are no NGS tests approved for clinical use – existing commercial tests are for research use only and clinical trials.

Malaria is caused by eukaryotic *Plasmodium* parasites and is commonly diagnosed using rapid diagnostic tests that detect parasite antigens from a sample of blood. Sequencing is most commonly used in the research setting, for example to study the evolution of the parasite and the development of drug resistance. As sequencing technologies improve and reduce in cost, areas identified where near patient sequencing could have value include in supporting drug resistance surveillance, vaccine development, and in the field transmission and outbreak surveillance.

Key themes

A number of factors affect the implementation of NPPS, which can be particularly acute in lower resource and/or remote near patient settings.

Availability of equipment and consumables. Procurement challenges in terms of relative costs of equipment and consumables, and in obtaining these, are an ongoing issue for many lower- and middle-income countries. Efforts to mitigate these challenges are required, along with development of more streamlined workflows requiring fewer steps, automation of steps and processes, and lower volumes of consumables.

Availability of expertise. A range of expertise is required to deliver sequencing services and the poor availability of bioinformatics expertise in lower-resource settings is a particularly acute challenge. Increasing expertise can be done via a range of approaches, including 'train the trainer' models and the development of more automated workflows which will decrease the amount of training required.

Infrastructure and logistics. Given the number of possible NPPS approaches there is also wide variation in the infrastructural resources required, e.g. electricity, water supply, computational power, internet access and additional laboratory services such as culture. While culture-free sequencing methods might become more prevalent in the future, culture is still a requirement for the accurate diagnosis of many pathogens and can also provide samples for sequencing.

Data analysis and management. There are many considerations around the capture of NPPS data and how it is analysed, stored, reported, and shared. Those particularly relevant to the implementation of NPPS are that it is likely to be delivered in resource limited environments, therefore mixed models of proprietary and open source tools, offline analysis, and cloud-based services analysis and storage services are likely to be necessary.

Conclusions

Near patient pathogen sequencing has utility in genomic surveillance during disease outbreaks and could in the future inform the development of more rapid diagnostics of disease or conditions such as respiratory infections.

In some cases, NPPS is being deployed via mobile laboratories in the absence of established centralised sequencing facilities, or in situations where sequencing facilities are too geographically distant to support timely genomic epidemiology. There is, therefore, a question around how extensive NPPS facilities would need to be, should more reliable and centralised 'hub' sequencing laboratories be established in countries where they are currently few in number.

Many of the features of sequencing platforms that make them suitable for use in a near patient setting also benefit delivery of sequencing in lower resource settings more broadly – for example automation, lower consumable costs, and open-source analysis software. Near patient sequencing approaches can therefore provide useful examples of what can be achieved, such as innovation in mobile and field laboratories which demonstrates that sequencing can be delivered under extreme conditions, with limited available resources.

The use of NPPS in mobile and field laboratories also raises considerations around longer-term capacity building. These laboratories have, in some cases, bridged the gap between emergency and more established services. There are considerations around how near-patient approaches such as these could be embedded as more established services, for example to help manage seasonal disease outbreaks. New services could be established from scratch, or mobile laboratories could augment existing centralised laboratory services and be dispatched where needed. During the 2013-16 Ebola epidemic, and during the current (as of June 2022) COVID-19 pandemic, there has been much capacity building for NPPS facilities and efforts will be needed to ensure that this is not lost. The opportunity to

repurpose these resources to cover endemic diseases, surveillance and local public health priorities should be considered. These could include AMR and OneHealth surveillance. While efforts in this area are underway, establishing services such as these will require reliable long-term funding to ensure stability and ongoing service provision.

The role and value of NPPS has been demonstrated to a certain degree but considerable further development is needed to make it a routine option for health authorities in low resource settings. This development will require close and effective collaboration between companies, researchers, global health agencies and most importantly the healthcare professionals and public health authorities who will be commissioning and using these facilities in the future.

1 Introduction

Pathogen genome sequencing is extensively used in the management of infectious diseases, supporting diagnosis, treatment decisions, outbreak management and surveillance. The majority of this sequencing takes place in centralised laboratories, with samples being transported away from the patient setting for sequencing.

Recent infectious disease outbreaks, including the 2013-2016 Ebola epidemic in West Africa and the ongoing COVID-19 pandemic, has led to considerations around more flexible, near patient approaches to the delivery of laboratory services, including sequencing. For example, during the Ebola epidemic, the European Mobile Lab project established a number of field laboratories in West Africa, significantly reducing diagnostic turnaround times [1]. Researchers from the University of Birmingham, UK, deployed a 'lab in a suitcase' containing portable sequencing technology that was used to support on the ground genomic epidemiology in Guinea [2].

Near patient testing (NPT) of pathogens – including near patient pathogen sequencing (NPPS) – speeds up the time to results by removing the requirement to transport samples elsewhere and by allowing for lower throughput sequencing that can provide more rapid results. Quicker access to sequencing results enables close to real-time tracking of disease outbreaks, transmission, and identification of new pathogens or variants. Furthermore, it allows for the rapid diagnosis of patients and the potential to identify any susceptibility or resistance genes associated with treatment response to enable timely and informed decisions on patient care.

However, there are also several considerations associated with near patient pathogen sequencing, particularly in lower resource settings, to enable adoption of this approach, including:

- Infrastructure – facilities may be portable, temporary, or permanent. Resources may vary depending on setting, with the availability of space and access to internet and electricity being key considerations that may impact on what is possible and what is required
- Sequencing technology – the infrastructure available will impact on the sequencing technology that can be used, with some technologies being more amenable to the near patient setting than others
- Methods used – this includes the sequencing method, which will depend on what the requirements are for sequencing and the available technology but may involve metagenomics, panel testing (pre-defined genetic targets in a physical test) or whole genome sequencing (WGS), as well as prior sample processing, including any requirement for culture or amplification, and library preparation
- Data analysis and storage – whether this will be undertaken locally or if there is a requirement for cloud storage or analysis, which may depend on the availability of simple analysis pipelines or personnel with expertise in this area. There are also considerations around how widely available the data will be made and in what timeframe
- Cost and affordability – both in set up and for continued use
- Regulatory and standardisation requirements – to ensure approaches are beneficial and standardised so comparisons can be made across different sites.

FIND, the global alliance for diagnostics, wishes to understand the opportunities and challenges of near patient pathogen genome sequencing and support its implementation, with a view to providing more flexible and timely solutions to infectious disease management.

1.1 Near patient testing and sequencing

Near patient testing devices are defined in the EU IVD regulations as: “any device that is not intended for self-testing but is intended to perform testing outside of a laboratory environment, generally near to, or at the side of, the patient by a health professional” [3]. The evidence review outlined in this report uses the EU IVD definition of NPT devices as a basis for understanding the uses of pathogen NPPS, including the technologies used to deliver it.

There are many circumstances in which pathogen NPPS could be used – therefore, for the purposes of this report, the ‘laboratory environment’ excluded in the EU IVD definition will only be centralised sequencing laboratories where patient samples have to be transported away from the treatment centre or sample collection site, as opposed to local or temporary laboratory settings with the equipment required for sequencing. These settings could be permanent, semi-permanent or fully portable, located within hospitals, clinics, other healthcare settings, or in the field.

The term point of care testing (POCT) can be conflated with NPT, although POCT typically refers to testing conducted at the bedside with strict timeframe limitations, typically less than an hour, that would exclude NPPS approaches. In this report, near patient will be used as a broad term to include, but not be limited to, POCT.

There are currently no clearly defined timeframe limitations for near patient sequencing, therefore a flexible approach will be taken in this report with cases being assessed individually as to whether they are classifiable as near patient. This will be based on sequencing taking place in locations geographically close to patients, outside of reference laboratories, and utilising rapid sequencing approaches. The timeframe requirement for results to be useful and acted upon will also be taken into consideration, as this will vary depending on the specific clinical or public health scenario being addressed.

This report outlines the current near patient pathogen sequencing landscape and identifies the opportunities and challenges associated with its implementation.

2 Near patient testing and sequencing

In this chapter, we first outline current NPT approaches for infectious disease management and the advantages and limitations of these, before discussing the principles of NPPS and its advantages and limitations over current non-sequencing NPT approaches.

2.1 What is near patient testing?

Near patient testing is any testing that is done near to the patient, outside of traditional centralised or reference laboratory settings, that reduces the time to results, and supports more rapid and informed decision making. This is due a number of factors including reduced sample transport times, the ability to carry out lower-throughput testing, and the use of more rapid technologies. There are two broad areas where this can have a significant impact:

- Disease epidemiology (population level): outbreak management and surveillance.
- Clinical (individual/patient level): aiding with diagnosis and treatment decisions.

These two areas are interlinked, with testing used for patient diagnosis also allowing for the study of disease epidemiology, as well as identification of novel or emerging pathogens, depending on the method used and availability of the resulting data.

Furthermore, there are two main settings where NPT could have an impact, with these having different infrastructure capabilities and requirements:

- In the field – likely to require portable testing capabilities that can be transported to outbreak areas. NPT could be useful for identification of novel pathogens or variants, outbreak management, and aiding with diagnosis
- In healthcare settings (including hospitals and community health clinics) – likely to have permanent, dedicated laboratory space, although this may be limited. NPT could be useful in diagnosis and treatment decisions, as well as managing and tracking nosocomial infections.

There may also be benefit in utilising NPT at points of international travel, such as at airports and harbours to prevent global spread, particularly of novel pathogens [4].

2.2 Characteristics of laboratory tests

There are a number of operating characteristics for laboratory tests, these include [5]:

1. **Reliability**. This is a measure of reproducibility observed by running the test many times on the same specimen. Reliable tests can be inaccurate, but consistent in returning inaccurate results. Changes in the results of a reliable test likely mean a true change in a patient's condition
2. **Accuracy**. This reflects the extent to which the test result represents the 'true' value of the parameter being measured. Accuracy and reliability can change independently of each other
3. **Sensitivity**. The proportion of patients with a disease who have a positive diagnostic test
4. **Specificity**. The proportion of patients without a disease who have a negative diagnostic test
5. **Predictive value**. The likelihood that a patient does or does not have a disease, given a positive or negative test result.

Test purpose has an impact on the characteristics of the type of test you use.

- **Diagnosis**. The ideal diagnostic test has high sensitivity and specificity – however in practice very few tests meet these criteria. For ruling out disease, it is usually more appropriate to use a test with high sensitivity to minimise false negative results. To confirm disease, tests with high specificity will minimise false positives. In most

situations, test purpose will inform the trade-off between sensitivity and specificity in terms of choosing the most appropriate test
- **Monitoring**. Test reliability and accuracy are more important considerations since this use requires making multiple measurements over time, which need to be directly comparable
- **Research**. Another category where laboratory testing can be useful is in research, for example to understand the pathophysiology of disease or other disease process.

The laboratory testing cycle has three main phases: pre-analytical, analytical, and post-analytical [6, 7]. There are a number of considerations and potential sources of error that could affect test performance during each of these phases. Most errors occur in the pre- and post-analytical phases (error rates 46-68% pre-analytical, 7-13% analytical, 19-47% post-analytical) [8].

Pre-analytical phase. The decision is made to perform a test and it is ordered, the sample is taken and is transferred to the laboratory. Processes that occur during this phase include patient and test identification, collection of information about the patient and any relevant data entry, sample collection and pre-processing, and sample transport to the laboratory. Errors can occur during any of these processes; common errors include ordering the wrong test, using an inappropriate container for the sample, and sample mislabelling.

Analytical. During this phase, the test is carried out. This phase requires standardisation, quality assurance processes and trained laboratory personnel. Sources of error can occur if laboratory equipment malfunctions, there is a failure of quality control procedures, or procedures are not followed correctly. There have been long term and ongoing efforts to reduce errors in this phase, via QC programmes and other operating procedures and guidance. As such, analytical errors for most tests are declining long term and are less likely than pre- and post-analytical errors [8].

Post-analytical phase. Once the test has been carried out, a report is generated and the result is conveyed to the clinician; the data are then interpreted which informs the clinical response. Data interpretation and analysis could be carried out by the laboratory and/or the clinician. Errors can include mistakes in the laboratory report or misinterpretation of results by the clinician.

2.3 Approaches to delivering testing services

Centralised or more decentralised approaches can be taken to delivering testing services. Centralised models typically consist of several hubs and a centralised reference laboratory. In this model, hubs collect all the samples within an area and may perform some sample processing before sending them on to a central reference laboratory, where all testing is carried out. Decentralised models, on the other hand, involve district level laboratories and regional reference laboratories that carry out testing for each area, with a central reference laboratory taking samples that cannot be tested elsewhere or for highly specialist testing. Decentralised models often also involve some testing being carried out at the health centre [9, 10].

While both of these models have proven successful in some instances, in settings where transport of samples to regional or centralised laboratories may pose serious challenges and be associated with significant delays, near patient approaches may greatly reduce the time it takes to get results. Key advantages and limitations of near patient approaches are summarised in Table 1.

Table 1. Advantages and limitations associated with near patient approaches.

Advantages	Limitations
Reduced time to results – quicker clinical decision making	Accuracy of the tests – some near patient approaches may not be as sensitive and specific as traditional centralised approaches; if tests are viewed as unreliable, they may not be widely adopted [4]
More flexible and/or portable testing	Standardisation – required to ensure consistent performance but may be challenging to implement
Results in real-time – increased detection and management of outbreaks	Increased workload for healthcare professionals at local healthcare sites
Reduced number of patients lost to follow-up – particularly in LMICs where there may be significant travel to local healthcare centres	Expertise to carry out tests is required – some near patient tests are simple and easy to use and interpret but others require more expertise
Testing prioritisation beyond first-come-first-served – prioritise patients who most need a rapid result	Increased procurement challenges when supplying consumables to many peripheral sites, as opposed to a single centralised laboratory
Increased reliability of the system – e.g. reduced chance of delays or sample loss due to transportation; less chance of sample deterioration which could contribute to potential test failure	

Near patient testing approaches may be a particularly good solution following natural disasters, after which infectious diseases often increase in incidence. In these settings, any established laboratories in the area are likely to have been affected and transport routes out of the area will be impacted. The ability to bring in mobile NPT solutions to the affected area to diagnose and monitor infectious disease may help to reduce transmission and improve treatment decision making.

2.4 What types of tools and methods are available for NPT?

Near patient pathogen testing can use a number of different methods, including antigen or antibody immunological tests, and polymerase chain reaction (PCR) tests. These can detect pathogens either directly, via direct detection of the pathogen (culture/microscopy), pathogen peptides (antigen testing) or pathogen nucleic acids, or indirectly, via detection of a host response (e.g. presence of antibodies). Direct testing is typically preferable at an early symptomatic phase of infection when pathogen load is high, while indirect methods may be better later, as antibody production does not happen immediately. This makes antibody testing more suited to pathogen surveillance, while direct methods are generally preferable for diagnostics. Advantages and limitations of some of these detection approaches are summarised in Table 2.

Table 2. Advantages and limitations of different pathogen detection approaches.

Test	Advantages	Limitations
Direct pathogen visualisation (culture or microscopy)	Not affected by genetic variation Can detect multiple pathogens Can determine pathogen load Culture allows for further characterisation and study of the pathogen, including antimicrobial susceptibility testing Microscopy is relatively cheap and easy to use	If using live samples may require high level biocontainment facilities Requires expertise and specific materials so often only conducted in centralised laboratories Culture can take considerable time and is not possible for some pathogens; only detects living organisms Microscopy cannot detect viruses or differentiate strains
Immunological tests – tests for the presence of antigens or antibodies	Can be highly sensitive and specific, even down to detection of different strains Low susceptibility to contamination Can be designed for rapid and easy detection, including for self-testing For rapid diagnostic tests reagents can be stored at room temperature	Can be prone to cross-reactivity to similar pathogens Different strains may have different reactivity Pathogens can mutate to a point where no longer detected Antigens and antibodies are detectable at different time points
Molecular diagnostics – detection of pathogen nucleic acids (excluding sequencing)	High sensitivity and specificity Rapid results possible Can be used to quantify pathogen load	Requires prior knowledge of pathogen genome Specificity can be impacted if pathogens mutate in a primer binding site Primers may have to be continually reviewed (time and expertise required) Access to the correct primers is required, with multiple sets sometimes required depending on pathogen and strain Can be issues with contamination Sensitivity can vary, particularly early in disease

There are several technologies that can be used to measure different pathogen biomarkers, which vary in complexity both in terms of the design of the test and in the equipment and expertise required to use them. There are a number of POC technologies that are particularly amenable to pathogen detection in near patient settings as they meet the World Health Organization (WHO) ASSURED criteria for POCT, namely that they are affordable, sensitive, specific, user-friendly, rapid and robust, equipment-free, and deliverable to end-users [11, 12]. These include [4, 11, 13-15] (Table 3):

- Lateral flow assays, in which small sample volumes move along a paper- or membrane-based platform before the pathogen gets captured at a positive line, with a control line indicating the test worked
- Microfluidics, which involves temporal and spatial control of very small volumes of fluid to give a result
- Plasmonics, which use metals (commonly gold, silver, or aluminium) and light to detect pathogens

- Nucleic acid amplification approaches, mostly PCR-based, including quantitative or real-time PCR, Loop-mediated isothermal amplification (LAMP), and quantitative nucleic acid sequence-based amplification (QT-NASBA).

These technologies can also be integrated, for example lateral flow assays can be based on microfluidics, microfluidics can be used for nucleic acid amplification, and plasmonics can be developed as a lateral flow assay [11, 16, 17]. Some technologies are more suited to detection of certain pathogen biomarkers, for example lateral flow assays are typically used to detect antibodies or antigens, plasmonics typically detect antigens, and nucleic acid amplification detects pathogen DNA or RNA. Although these tests are designed to be simple to perform and interpreted by the end user, development can be complex, meaning there may be delays to production in response to novel pathogens [18].

Table 3. Advantages and limitations of different POC technologies.

Technology	Advantages	Limitations
Lateral flow assays [4, 13]	Rapid, low cost, easy to interpret, low sample volume Single step assay that does not require any laboratory equipment, with no or minimal sample preparation Do not require a cold chain and typically have long shelf lives	Low sensitivity – a low pathogen load will often produce a false negative result Not quantitative Viscous samples can take longer to produce a result
Microfluidics [11]	Only requires small sample volume Automated and can give simple, easy to interpret results Paper-based is low cost and user friendly Textile-based tests are being researched which are low cost, potentially less complex to produce, and could be developed into wearables [13]	Complex development and standardisation Can have issues with sensitivity, depending on approach
Plasmonics [11]	High sensitivity and specificity Can use unprocessed samples (whole blood) Paper-based is low cost, only requires small sample volume, and is portable	Can be slower than other technologies (>1 hour) Complex development and standardisation
Nucleic acid amplification [13-15]	High sensitivity due to amplification Devices have been developed for POC that integrate preparation, amplification, and detection	Typically require more equipment than other tests, including requirement for continuous electricity Requires nucleic acid extraction step and may require thermocycling, which is energy intensive Accuracy can be reduced by presence of inhibitors in unprocessed samples

Although these POC tests can give cheap and rapid results, there are two major drawbacks when used diagnostically, namely the limited availability of multiplex tests that can detect multiple pathogens at the same time, although this is an active area of research, and the inability to detect novel pathogens. Other approaches may be possible when extending to a near patient approach, rather than POC, including culture and microscopy, that can identify multiple pathogens, although these require expertise in the technique and pathogen to allow for identification.

However, specific pathogen identification may not always be required, with tests that simply distinguish between bacterial and viral infections still having clinical utility when making treatment decisions and having the potential to reduce over treatment with antibiotics [19]. This clearly highlights that the specific use case greatly impacts on which test is best suited

and most effective, with different scenarios requiring different approaches to achieve optimum outcomes.

2.5 Near patient pathogen sequencing

The previously discussed NPT approaches can give rapid, easy to interpret results at low cost and with minimal expertise required for use, while also being highly portable. These NPT approaches are generally used for diagnosing individuals and have more limited capability in outbreak investigation due to their limited multiplexing and ability to detect novel pathogens or lineages. In contrast, NPPS allows for the sequencing of pathogens in near-real time, with the potential to give rapid answers to questions about outbreak origin and transmission, supporting treatment selection, and allowing for the detection of treatment resistance, novel pathogens or pathogen variants. The following sequencing approaches can be taken:

- Whole genome sequencing – captures the entire genome sequence of an organism, most often from cultured pathogen isolates
- Targeted panel tests – specific sections only of the genome are sequenced as a predefined test to capture data on particular genetic loci of interest
- Metagenomics – captures all genomic information in a sample without prior culture typically from a range of microbes

Alongside surveillance and monitoring, NPPS can also aid in diagnosing individuals and informing treatment decisions, with advantages over NPT generally being for patients with non-specific symptoms who would otherwise require multiple tests for different pathogens. However, compared to non-sequencing NPT, NPPS is associated with higher cost, decreased portability, and requires significantly more expertise both to run and analyse the results.

Some notable advantages of NPPS over simpler NPT can include:

- The ability to track pathogen lineages, giving information on routes of transmission, as well as allowing for the tracing of outbreak origin. Different pathogens have differing genome biology that means the utility of sequencing for surveillance varies, with some pathogens requiring more sequencing to be carried out to provide usable results [20]
- Not requiring any prior knowledge of the pathogen or target – applicable to some NGS approaches
- The ability to detect pathogen mutations associated with new variants and/or drug resistance or susceptibility
- The ability to get an overview of the whole genome and/or of multiple important areas of the genome if using targeted sequencing
- Pathogen sequencing early on in an outbreak can enable the development of simpler and more rapid diagnostic tests and inform vaccine design

These advantages are not necessarily applicable to all NGS methods, for example some approaches require prior knowledge of the pathogen – the benefits and limitations of different sequencing approaches, in different near patient contexts, are discussed in detail later in this report.

Although NPPS has some advantages over the NPT approaches outlined above due to the extra level of detail that can be gained, a number of limitations have already been highlighted and these are analysed in more detail in the following chapters. These include:

- **Infrastructure**. Needs vary according to the technology but access to reliable electricity, cold storage, climate control and sufficient physical space for the equipment required for preparation and sequencing should be considered

- **Availability of pathogen DNA or RNA.** This can be limited, depending on pathogen load across the disease course and on use of prior treatment, for example initial treatment with antibiotics
- **Expert personnel.** Bioinformaticians are required due to complex data analysis needs, and trained laboratory staff may also be necessary for sample preparation and to run the sequencing
- **Data analysis and storage**. Sequencing data requires more complex analyses and increased data storage capacity, particularly in the case of whole genome sequencing (WGS) [21]
- **Cost**. Sequencing capability is more expensive to set up and run than other NPT approaches.

The next chapter explores the applicability of different sequencing technologies to near patient contexts in more detail.

3 Near patient pathogen sequencing technologies

Sequencing is the process by which nucleic acids in a sample are converted into data that can then be analysed. There are different ways that this process can be achieved and a number of different technologies that can facilitate it. These techniques include:

- Nanopore sequencing (Oxford Nanopore Technologies)
- Sequencing by synthesis (Illumina)
- Ion torrent semiconductor sequencing (ThermoFisher Scientific)
- DNA nanoball sequencing (Beijing Genomics Institute and MGI Tech)
- Single molecule real-time sequencing (Pacific Biosciences)

Each of these techniques, which are broadly aligned to specific companies, can be performed on a range of instruments produced by the developers named above. This is an active area of commercial research and development, and it is likely that further sequencing platforms from existing and new providers will become available in the future.

Most of these technologies have been developed for high-throughput sequencing, methods which allow large quantities of DNA or RNA to be sequenced, and reduce the relative costs of sequencing per sample [22]. However, access to these technologies remains limited with a complex laboratory ecosystem required to utilise these sequencing platforms, for example those required for BGI/MGI and Pacific Biosciences [23]. Contributing factors include infrastructure requirements, limitations on supply chains for reagents, staff training requirements, longer laboratory turnaround times, bioinformatics processing and costs of the sequencers [24]. These factors can be barriers to use of high-throughput sequencing for NPPS.

A combination of research and stakeholder interviews were used to identify sequencing technologies being used for NPPS. In this section, only technologies that have been applied in the near-patient context or in a field laboratory setting will be described in detail, although this landscape is likely to change in the future. We also present the ideal characteristics of near-patient sequencing platforms, the current NPPS platform landscape, and consider future innovations in the sequencing market.

3.1 Choice of sequencing platform

Choice of sequencing technology in a near patient setting will be governed by a range of factors in addition to the requirements for sequencing in a more centralised or reference laboratory setting. For example, the World Health Organization's GLASS report outlines the infrastructure requirements for high-throughput sequencing [25]. These include:

- Reliable Internet connection
- Continuous A/C electricity supply
- Vibration-free platforms
- Dust control for some equipment
- Molecular biology-quality water (which can be purchased in bottles or produced on-site by treatment or filtration)
- Temperature and humidity regulation within an adequate range for chemical reactions and temperature-sensitive equipment
- Cooling and storage for reagents and DNA, with periodic recording of storage temperatures
- Rooms that are air-tight and maintain a stable operating environment
- Enough space to set up a unidirectional workflow
- Automated liquid handling systems or robots for high-throughput capacity.

In most circumstances, these requirements are more challenging to deliver in near patient settings and will impact on choice of sequencing technology. Overarching considerations will

depend upon the sequencing objectives, which involve a trade-off between accuracy, efficiency, time to result and cost. Factors influencing the decision to use NPPS will include: the status of existing sequencing infrastructure and centralisation; availability of diagnostics; availability of clinical microbiology laboratories; complexity of the infectious disease and healthcare system landscape; surveillance infrastructure; and importance for management of outbreaks of unknown origin. Evaluation of the use of sequencing and the cost-benefit implications will need to be performed on a case-by-case basis.

More detailed decision making around choice of sequencing platform will consider:

- Where sequencing is being deployed – e.g. in a mobile laboratory or static site – and minimum environmental requirements (see Chapter 5)
- Size, weight and power requirements and how easy it is to move the technology, should that be required
- Technical characteristics of the platform:
    - Throughput levels, including minimum samples and pooling/barcoding possibility (sample capacity per run and per kit size)
    - Yield and quality per sample
    - Run time and time for analysis
    - Type of sequencing – chemistry and read length considerations
    - Scalability, flexibility and connectivity
- Costs related to equipment (sequencer and associated equipment), establishment of workflow, reagents, other consumables, personnel time, bioinformatics resources, analysis, and maintenance. Costs will occur both in set up and on an ongoing basis – which can be evaluated as cost per sequencing run, per genome sequenced, or per megabase of output data
- Availability of reagents and consumables with associated procurement pipelines, including transport and appropriate storage
- Established personnel and support for troubleshooting, protocols, and pipelines for analysis.

3.2 Oxford Nanopore Technologies

Oxford Nanopore Technologies (ONT) produces a range of sequencing systems based on nanopores. ONT's systems are designed to be relatively mobile, generate ultra-long reads and be more accessible to those with less experience and expertise. The systems are relatively low cost and are provided primarily through equipment purchase and customer support. Additionally, equipment can be rented from ONT.

Extracted nucleic acids are prepared for sequencing by ligation of a motor protein and adapter sequence at the ends of each strand. RNA may either be amplified and converted to cDNA prior to sequencing as occurs with other sequencing systems or the RNA can be read directly (direct RNA sequencing) without prior amplification or conversion to cDNA.

During sequencing, tagged, single stranded DNA or RNA molecules are fed through a membrane-bound protein pore – a 'nanopore' – by a motor protein. As each DNA or RNA nucleotide is fed through the nanopore, it interrupts the electrical current that exists across the pore and these signals are detected by the sequencing system. This pattern of disruption can be read to determine the base sequence of the molecule.

ONT systems are capable of producing very long sequencing reads, up to around 800,000 bases in length. Sequencing is also fairly rapid: one nucleic acid can be read by each pore at a given moment, and each molecule is read at approximately 400 bases per second. ONT sequencing systems provide different capacity, throughput and mobility, and cover a wide range of price points.

Oxford Nanopore Technologies has recently established Oxford Nanopore Diagnostics to focus on the future potential diagnostic use of nanopore sequencing.

Table 4. Information on ONT sequencing platforms

Sourced from nanoporetech.com

**A Starter Pack provides users with the opportunity of accessing the technology by using consumable budget. for lower cost, however, the user would not own the device. Users may purchase the device for additional cost.*

	MinION Mk1B	**MinION Mk1C**	**GridION Mk1**	**PromethION 24**	**PromethION 48**
Run time – flexible dependent upon data required	1 min - 48 hrs	1 min - 48 hrs	1 min - 48 hrs	1 min - 72 hrs	1 min - 72 hrs
Maximum output	50 Gb	50 Gb	250 Gb	5.2 Tb	10.5 Tb
Size (cm) H x W x D	2.3 x 10.5 x 3.3	3 x 14 x 11.6	22 x 36.5 x 36	19 x 59 x 43	19 x 59 x 43
Weight (kg)	0.087	0.45	11	28	28
Power requirement	Powered by laptop.	25W	800W	2kW	2kW
System price starter pack*	$1,000	$4,900	$49,995	$225,000	$310,000
Read length	Dependent on length of target molecule maximum to date > 4Mb.				
Description	Low cost, mobile, long read sequencer.	Mobile long read sequencer and analysis platform in one unit	Medium capacity desktop long read sequencer with onboard analysis platform	High capacity desktop long read sequencer	Higher cost, high capacity desktop long read sequencer

Advantages of Oxford Nanopore Technologies sequencing

- Rapid and flexible, particularly useful for sequencing smaller genomes. Sequences read in real time, allowing for termination when user determines enough reads have been generated
- Smaller sequencing units can be purchased at low cost
- Mobile sequencing is possible – the small size and high portability of some systems means that these can be used in the field
- Some reagents use dry lyophilisation technology and do not require cold storage meaning they can be safely stored for a set period in environments where refrigeration is not possible or unreliable
- A simple user interface and analysis platforms for base calling and analysis are under development
- Direct RNA sequencing and methylation sequencing are possible

- Automated sample preparation technology available.
- Relatively inexpensive at low throughput.

Limitations

- More limited barcoding of samples limits multiplexing for some applications with higher costs per sample for high-throughput sequencing applications
- Some techniques use reagents requiring cold storage meaning these approaches can only be used with reliable cold storage and reagents may have a short shelf-life (<3 months)
- Dry lyophilised kits require intact storage bags in order to maintain low humidity
- Raw signal output files are very large which has an impact on data storage, data availability for analysis and subsequent data deposition into databases
- Bioinformatics expertise often needed to develop analysis pipelines, although examples of use by non-experts once available
- Lower read coverage than short-read platforms
- Higher base-call error rate and overall error rate
- Current automation limited in terms of sample barcoding and added cost of sequencing.

Appendix 8.2 contains a summary of the advantages and limitations of different sequencing platforms.

3.2.1 MinION Mk1B and MinION Mk1C

The MinION devices are portable sequencers that enable real-time DNA or RNA sequencing [26]. There are two MinION devices available and the key difference relates to the way these devices are powered. The MinION Mk1B is a USB-sized device which can be powered from a laptop or a MinIT, a companion device to the MinION pre-configured with the software to carry out data acquisitions and perform base calling. The MinION Mk1C is an all-in-one device with a touchscreen, simple device control and visualisation of results, which facilitates integrated computing for base-calling and analysis software. MinIONs can generate the same volume of genomic data, using the same reagents [26]. These platforms can be used with the Flongle flow cell to generate 2.8 gigabases (Gb) or the MinION Flow cell to generate 50Gb.

The Flongle is a flow cell adapter that can be used with the MinION or GridION to enable generation of sequencing data at the lowest overall cost on ONT sequencing platforms [27]. It can sequence individual samples or use barcoding for multiplex sequencing. Applications include amplicons, panels or targeted sequencing, quality testing and small sequencing tests. It only produces a low volume of data compared to other flow cells making the utility of this device limited compared to the MinION flow cell. However, Flongle flow cells are cheaper than MinION flow cells and, as a result, it may be advantageous to deploy the Flongle as part of a rapid sequencing response at lower initial costs. The lower yield of data make it unlikely that the Flongle would be used where the purpose requires higher data resolution, e.g. transmission dynamics or identifying resistance genes [28].

The MinION devices offer many of the same advantages as other NGS platforms – for example they support multiplexing of samples – but are also portable and inexpensive relative to other platforms, allowing for in-field deployment [29]. While they are designed to be used between 18°C to 24°C, they may be adapted to work at higher or lower temperatures, although one interviewee indicated that MinION performance may be impacted by extreme high temperatures (>30°C). MinIONs have been deployed to perform real-time, field sequencing in a range of environments, including space [30], Antarctica [31], rainforest [32], desert [33], and ocean [34]. Researchers have also demonstrated the ability to perform sequencing using only solar power as an energy source [35].

ONT protocols can be completed quickly relative to those for other platforms with a number of studies demonstrating generation of results within 12 hours of sample collection [36-39]. One study assessed the use of MinION to perform real-time sequencing of primary respiratory samples from *M. tuberculosis*-infected patients. This study demonstrated that it is possible to carry out pathogen identification, phylogenetic placement and initial susceptibility predictions after 7.5 hours and generate complete results after 12.5 hours using the MinION [39]. This run-time includes decontamination, DNA extraction and sample preparation.

A number of studies have demonstrated use of MinION offline and without the need for an internet connection, significant in low-resource settings where internet connection may be unreliable [36-38, 40]. However, it may not be possible to perform all types of analysis, depending upon the nature of the tool or database required.

It has been reported in the literature that ONT has an overall lower sequencing quality in terms of the accuracy of base calling than other sequencers, including Illumina, which is the platform used most frequently for comparison. There have been significant improvements in base calling quality from ONT but these improvements can be difficult to evaluate from the literature because there is still variation in how this quality is reported [41, 42]. ONT has demonstrable use for a range of purposes with growing evidence that the base calling quality is sufficient for these applications, although it is not yet clear if ONT quality would be considered sufficient for clinical diagnostic use (e.g. Chapter 6). This means that the quality of sequencing data may limit analysis and interpretation of results for certain purposes [43]. The importance of data quality will need to be balanced against more practical considerations underpinning the choice of sequencing technology, particularly ease-of-use and portability. Additionally, changes in the technical performance of ONT platforms, composition of reagents and evolution of bioinformatic tools make it challenging to develop a consistent start-to-finish protocol [44]. Ongoing research and development efforts by ONT are beginning to address this, particularly through the work of Oxford Nanopore Diagnostics.

3.2.2 Automation: VolTRAX

The VolTRAX is an automated and integrated sample preparation platform for nanopore analyses [45]. It is channel-less and instead works by moving the sample over each component for thermocycling or incubation. Once sample preparation is completed, it only requires pipetting to the sequencer by the user. The VolTRAX can be considered an emerging product since it is currently research use only, however it has established proof-of-principle that automation is feasible for ONT devices. Low sample multiplexing – currently restricted to 10 samples – is the main reason for the relative high cost associated with sequencing using VolTRAX. Reductions in cost achieved through reduced costs of reagents, consumables and improved multiplexing will make VolTRAX more accessible in the future for near-patient sequencing applications.

3.2.3 Sequencing platforms in development

Oxford Nanopore Technologies have a number of sequencing platforms in development:

SmidgION. This is a smaller device which can be used with a smartphone or other mobile low power device for field-based analyses [46]. Potential applications including remote monitoring of pathogens in an outbreak and on-site analysis of environmental samples. This platform will be supported by rapid library preparation kits and ONT-provided analysis workflows.

MinION Mk1D. Increased processing capabilities of tablet devices mean that they are on the verge of being able to run a MinION sequencer and analyse the data. The Mk1D is designed to be an accessory keyboard with an integrated sequencer for tablet devices [47]. Users will be able to dock a recommended tablet, pair it and initiate sequencing.

PromethION P2 and P2 Solo. These platforms are designed as smaller and more compact forms of the PromethION 24 and PromethION 48 [48]. They are designed for lower throughput laboratories or where a laboratory wants to explore the use of this platform without committing to the larger machines. The PromethION P2 is a self-contained benchtop sequencer containing sufficient graphics processing units to run two high-output PromethION Flow Cells. The P2 Solo is a sequencing unit for two PromethION flow cells and can be powered using the GridION Mk1 or other suitable computers. This platform could have some utility for NPPS with higher throughput, a smaller footprint and lower upfront cost than previous PromethION sequencers. This platform is currently available for pre-order for research use only. The list price for the system starter pack for these two devices is US$10,455 and US$59,995 respectively.

3.3 Illumina

Sequencing using Illumina systems provides high throughput short-read sequencing by synthesis (SBS), which is widely used.

Sequencing by synthesis nucleotide identification occurs as modified nucleotides are incorporated into newly forming DNA. Fluorescently tagged (modified) bases are detected as they are incorporated. Unlike early chain-termination methods (Sanger sequencing), these fluorescently tagged bases do not cause DNA synthesis to stop. Each time a base is incorporated, the attached fluorescent tags are washed away after detection, allowing for more modified bases to be added after this point. The process is repeated until the maximum number of cycles (and therefore sequence length) is reached.

Illumina produces a range of platforms which cover a large array of sequencing applications; and these differ in size, capacity, and cost. Illumina's smaller sequencing platforms are described below (Table 5). In addition, Illumina provide the NextSeq 1000 & 2000 series and the NovaSeq 6000. These platforms are significantly larger and unlikely to be practical or economical outside high-capacity and well-resourced laboratories.

Illumina has announced recent improvements in sequencing chemistry, codenamed Chemistry X, which promises to deliver two times faster cycle times, two times length sequencing reads and three times greater accuracy [49]. These improvements have the potential to reduce the cost of sequencing. It is anticipated that this new sequencing chemistry will be incorporated into future sequencing platforms.

Illumina sequencers become cost-effective when processing high numbers of samples. However, the overall cost of each run is high, and optimisation of each run is necessary to ensure that they are cost-effective, reducing the flexibility of their use. Given the size, weight and environmental condition requirements of Illumina machines, only the Illumina iSeq 100 will be considered here for potential utility in the near patient setting.

Table 5. Information on Illumina sequencing platforms. *Note: list prices are not publicly available on the Illumina website and therefore have not been included here*.

Sourced from illumina.com

	iSeq 100	**MiniSeq**	**MiSeq series**	**NextSeq 550 series**
Run time	9.5–19 hr	4–24 hr	4–55 hr	12–30 hr
Maximum sequence data output	1.2 Gb	7.5 Gb	15 Gb	120 Gb
Maximum read length	2 × 150 bp	2 × 150 bp	2 × 300 bp	2 × 150 bp
Size (cm) H x W x D	42.5 x 30.5 x 33	82.5 x 86.4 x 63	68.6 x 52.3 x 56.5	58.5 x 53.4 x 63.5
Weight (kg)	16	59	57.2	83
Power requirements	80W	240W	400W	600W
Description	Lower cost, lower capacity, smallest form benchtop sequencer available from Illumina. Released 2018.	Lower cost, lower capacity benchtop sequencer. Released 2016.	Mid-range benchtop sequencer providing longest reads available on Illumina platforms. Released 2011.	Mid-range benchtop sequencer, providing greater throughput than cheaper devices. The related *NextSeq 550 Dx* is diagnostic tool for specific clinical applications only [50].

Advantages of Illumina sequencing

- Comparatively low-cost sequencing at high throughput, where many samples are being processed
- High raw read accuracy and read depth generating high accuracy data
- One of the more commonly used systems for high resolution genomic analysis allowing collaborative development of expertise and advances. Many genetic or research laboratories already possess these systems, and bioinformatics pipelines are relatively well-established
- High levels of sample multiplexing are possible, meaning a high number of samples can be run at once
- Option of targeted and WGS approaches.

Limitations

- Longer sequencing run time
- Most platforms are large and costly to purchase, some require specialised infrastructure for safe use
- Short reads limit accuracy in complex genomic regions and opportunity for identification of the genomic context of mobile genomic elements
- Installation of Illumina machines involves a number of environmental considerations, including vibration, placement requirements, and control of temperature, humidity, elevation and air quality.

Appendix 8.2 contains a summary of the advantages and limitations of different sequencing platforms.

3.3.1 *Illumina iSeq 100 applications*

The Illumina iSeq 100 is the cheapest and smallest in terms of size and weight (Table 5) of Illumina's NGS systems. The iSeq is comparatively simple to set up without a trained technician [51] and is more cost-effective than other Illumina systems when carrying out lower-throughput sequencing [52]. Illumina sequencers typically utilise four-channel or two-channel sequencing-by-synthesis (SBS) chemistry to detect incorporation of different nucleotides. The iSeq platform utilises a slightly different sequencing chemistry to other Illumina sequencing platforms based on complementary metal-oxide semiconductor (CMOS) technology and one channel SBS chemistry. The iSeq has been adapted for bacterial and viral sequencing and data analysis can be performed using the sequencer, including read alignment and variant analysis.

One study assessed the performance of Illumina sequencing of *M. tuberculosis* using two MiSeq instruments and one iSeq 100 instrument. This study compared both laboratory performance and wider practical considerations related to the setting and intended use. The iSeq presented advantages in terms of hands-on-time, lower capital costs and sequencing data processing within 24 hours [53]. Optimisation of sample runs on the MiSeq resulted in lower cost per sample than the iSeq at a higher throughput. The iSeq was more affordable than the MiSeq when processing up to 1000 samples at 50X coverage. This suggested that the iSeq100 may be a more cost-effective sequencing solution in situations where it is more difficult to predict the throughput of samples and ensure optimisation of the run. This study identified uniform coverage between the sequencing platforms evaluated with no difference between sequencers and high percentage agreement for variant calling. Other studies have also identified high concordance between the iSeq 100 and MiSeq [54].

Expert interviewees stated that implementation of Illumina sequencers in many lower resource settings remains operationally and logistically challenging, for example in situations where the electricity supply and/or internet connectivity are unreliable. There are also challenges around procurement of reagents and regular servicing of sequencing platforms or repairs, which can result in machines not being operational. One expert commented that the iSeq had proven significantly more usable than their existing Illumina machines given these types of challenges. For this sequencer, they however faced delays in support resulting in long periods where these machines could not be used. They highlighted the value of external support to implement a pre-curated, rapid bioinformatics pipeline when they did not have in-house bioinformatics expertise to achieve this. This has supported development of in-country expertise for sample collection, processing and NGS analysis [55]. Experts indicated that the iSeq will be most appropriate where there is some laboratory infrastructure, but that this could be more basic than is typically required for other Illumina sequencers.

The main evidence for use of the iSeq in a near patient setting comes from the Democratic Republic of Congo where this sequencer was deployed by the Institut National de Recherche Biomédicale (INRB) to support the outbreak response to Ebola within six weeks of the outbreak being declared [51]. North Kivu had a high burden of Ebola cases and therefore a field genomics laboratory was deployed to this region by the INRB [56]. While this improved turn-around-time between sample collection and sequencing, various infrastructural, logistical and funding challenges continued to impact speed and consistency of any sequencing performed.

3.4 Ion Torrent

ThermoFisher supplies semi-conductor sequencing through its Ion range of systems. Sequencing utilises a semi-conductor chip and a bead-based system which provide a physical platform for sequencing through DNA synthesis. This system relies on the detection of hydrogen ions released as nucleotides are incorporated.

The Ion GeneStudio S5 is the newest and highest capacity NGS system in this range. Older systems include the Ion Proton system and Ion Personal Genome Machine (PGM). An automated 'specimen to report' system, the Ion Torrent Genexus, provides rapid sequencing with reduced sample contact [57]. A number of sequencers in the Ion Torrent range are described below. Automation of sequencing sample preparation can be performed using the Ion Chef system to provides automated library preparation, template preparation, and chip loading reducing manual time [58].

Table 6. Information on Ion sequencing platforms. *Note: list prices are not publicly available on the ThermoFisher website and therefore have not been included here*.
Sourced from thermofisher.com

	Ion PGM system + Ion 318 chip	**Ion Proton system + PI chip**	**Ion GeneStudio S5 System + Ion 540 chip**	**Ion GeneStudio S5 Prime System + Ion 550 chips**	**Ion Genexus system**
Total sequencing and analysis time at max. throughput	7.2 hr (400bp)	4 hr	19 hr	6.5 hr	24 hours
Max. throughput/day	2Gb	10Gb	15 Gb	50 Gb	Unknown
Read length	400bp	200bp	200bp	200bp	Unknown
Size (cm) (W x D x H)	61 x 51 x 53	54.2 x 77.5 x 47.4	54.2 x 80.6 x 50.9	54.2 x 80.6 x 50.9	106 x 81 x 167
Weight (kg)	39	59	63.5	63.5	204
Description	Large benchtop sequencer providing the longest reads available from the Ion sequencing range. Older system with low throughput.	Fast, modest throughput benchtop sequencer.	Lowest capacity and throughput of the Ion GeneStudio range of sequencers. Low capacity in comparison to many other technologies. Small to medium profile desktop machine.	Fast and highest capacity Ion sequencing system.	Specimen to report automated system, claiming 10 minutes of 'hands-on' time. Permits variable throughput.

Ion Torrent has the advantage that these systems support automation with both the Ion GeneStudio S5 system and Ion Torrent Genexus system constituting the sequencing instrument and an automated sample preparation system. This minimises the number of errors introduced from manual handling of these machines.

There are examples of sequencing using Ion Torrent sequencers in field settings. In early 2015, during the Ebola outbreak in Sierra Leone, a group of researchers trialled the use of a Thermo Fisher Scientific Ion Chef and Ion Torrent PGM sequencer [59]. They tested performance of the sequencing platform in the UK before transporting it to Sierra Leone, where sequencing of *Zaire ebolavirus* (EBOV) samples occurred. This system was installed in a lined, air-conditioned tent with sequencing operational in one day with first data files transferred to the UK after less than five days. This was an unconventional use of this

sequencing technology under harsh conditions including high temperature, dust, high humidity, unreliable power supplies, and complicated reagent transport.

Advantages of Ion Torrent sequencing

- Comparatively short sequencing runs enable faster return of results
- Low substitution error rate
- Some systems facilitate a highly automated workflow for easy adoption and consistent application of sequencing
- Option of targeted and WGS approaches
- Longer individual reads
- Relatively inexpensive at low throughput.

Limitations

- Lower throughput in comparison to other NGS technologies, therefore comparatively expensive at high throughput.
- Environmental constraints relating to humidity, temperature, vibration and air quality.
- Higher sequencing error rate
- Higher cost per sample
- Shorter reads compared to other NGS technologies and is not able to perform paired-end sequencing
- Fewer bioinformatics tools built for Ion Torrent-generated data.

Appendix 8.2 contains a summary of the advantages and limitations of different sequencing platforms.

3.5 Comparison of sequencing technologies

Worldwide, Illumina is the dominant sequencing company making up at least 90% share of the sequencing market [23]. This success has been driven by the ability to generate high quality sequencing data alongside reduced cost through higher-throughput and multiplexing of samples. When discussing the establishment of in-country laboratories, interviewees highlighted the challenge of achieving sustained use of Illumina machines. A number of factors contributed to this, for example, reliable electricity supply, technical support and procurement of reagents. Use of Ion Torrent for NPPS appears to be minimal, however, this may change in the future. ONT's MinION has been adapted for use for near patient and field sequencing by a number of initiatives using different levels of infrastructure (Chapter 5).

Expert interviewees were clear that ONT MinION was the only sequencing platform considered for near patient or mobile sequencing, because they felt it was the only sequencer on the market that met their need in terms of portability, ease-of-use and overall cost. However, they highlighted a number of limitations around use of the MinION, including lower data quality and availability of standardised protocols.

Portability: ONT's MinION is the smallest sequencer on the market and has been used from a suitcase requiring minimal laboratory infrastructure [36]. One example of use was for surveillance of swine flu at an agriculture festival, where a laboratory was set up and operated by two people to produce a high quality multiplexed NGS library in seven hours and perform real-time analysis and public health interventions [37]. ONT's MinION has also been adapted for use from a lab in a suitcase [36, 60]. In comparison, portability of Illumina sequencing is more dependent upon the laboratory infrastructure required to maintain a suitable environment (Chapter 5). Experts indicated that the iSeq 100 is more robust in terms of functioning following movement or vibrations than other Illumina machines.

Turn-around-time: Illumina has a longer turn-around time than ONT devices, with an Illumina MiSeq versus iSeq 100 versus ONT MinION taking up to 48 hours, 28 hours and 12 hours respectively [39, 61]. The biggest contributor to the difference in sequencing turn-

around-time for the ONT MinION and the Illumina iSeq 100 or MiSeq is the difference in the sequencing run time [62]. However, ONT also has a shorter sample preparation time and it is possible to perform real-time sequencing on the ONT MinION enabling bioinformatics to be performed before the run has been completed. Bioinformatics can only be performed for Illumina generated data once the run has completed.

Ease of use: There are more published protocols available for Illumina sequencers with greater clarity on the performance of different workflows compared to ONT platforms. Illumina sequencing is performed more widely than ONT sequencing and some experts highlighted a preference for Illumina. A key reason given for this was familiarity with the technology and more reliable data generated, although the precise reason for this difference in performance was not clear. There is ongoing work to establish robust workflows to support implementation of ONT sequencing. In terms of usability, environmental requirements (e.g. around temperature and vibration) mean that current Illumina platforms are unlikely to be used for field sequencing [36]. Use in a more static location, for example, to support a hospital, could be an effective use of the Illumina iSeq 100, where a permanent laboratory could be established. Additionally, Illumina sequencing will typically require extensive human expertise to support all aspects of the workflow from library preparation to instrument support [63]. Less expertise appears to be required to operate ONT sequencing platforms, although opinions on the suggested level of expertise and length of training required differed between expert interviewees.

Applications: The MinION can be used for direct RNA sequencing, however currently this approach has a low sensitivity and it is not possible to multiplex [37]. Further work is required to optimise the analysis algorithms and Q-score for viral RNA sequencing to ensure consistent quality. Currently, to achieve high quality RNA sequencing, cDNA synthesis is required for all sequencing applications – particularly Illumina where direct RNA sequencing is not possible. Additionally, RNA degrades more rapidly than DNA into shorter reads, and in this case, Illumina will generate higher quality data [51]. Direct methylation sequencing can be performed on ONT sequencers. Methylation sequencing on Illumina sequencers is performed using bisulphite sequencing with additional laboratory and bioinformatics requirements compared to ONT sequencing.

Clinical application considerations: Illumina sequencing is cheaper per sample when fully optimised with higher multiplexing capacity. Overall ONT is relatively more costly per sample; one example showed that 12 samples on a MiSeq or between 3 to 5 samples on a MinION cost £96 versus £101 to £172 respectively. This is explained by the fact that the MinION generates less data per flow cell and has more limited barcoding restricting the number of samples per run [39]. However, ONT platforms are cheaper overall and can be used at lower throughput making these devices more flexible and adaptable to some near patient settings. This can be advantageous in low-resource settings where optimisation of runs and planning capacity for sequencing will be more challenging. Another advantage for lower resource settings is the ability to control the sequencing time on ONT devices, which allows the user to determine when sufficient data has been obtained. However, ONT has made progressive changes to reagents and flow cells over time resulting in changes of performance of this technology [60]. These changes make it challenging to match biochemistry to flow cells or performance of analysis when validating workflows. Illumina technologies are more standardised with established expertise in clinical settings to develop and validate these standardised workflows. Consistency will be critical to validation of these workflows for NPPS.

Data accessibility: Illumina technologies require bioinformatics expertise for analysis [36]. Illumina sequencing run-time is incompressible with data only obtained once sequencing has finalised, whereas ONT has real-time sequencing and analysis, which can be performed off-line [37]. This allows the user to choose how long to sequence enabling them to continue sequencing until sufficient coverage is obtained. This is advantageous where there is

variable levels of DNA, for example, providing a potential solution to the problem of variable amounts of *M. tuberculosis* DNA in patient samples [39].

Data characteristics: ONT is a long read sequencing technology able to read longer contiguous strands of between 10,000–100,0000 base pairs of DNA [64]. There are inherent advantages to producing longer sequence reads. As read lengths increase, it is more likely that a read will be distinct from other reads. This allows them to be computationally reassembled with less ambiguity. Long read data will have some advantages, for example when determining if resistance genes are located on mobile genetic elements.

Data quality: The literature around data quality for ONT MinION is variable although there is consensus that the quality and yield of data is lower than Illumina sequencing platforms. It is likely that, depending upon the purpose of use, the quality of data that is needed to achieve sufficient confidence in results will differ. For example, a comparison of ONT GridION and the Illumina MiSeq identified that the GridION provided lower coverage, high number of indels and was sensitive to SARS-CoV-2 viral load indicating the Illumina MiSeq is preferred for SARS-CoV-2 genomic surveillance [65]. MinION broadly is believed to yield sufficient quality data to reliably perform pathogen identification, with some exceptions [66]. In addition, it is possible to study transmission dynamics and identify resistance genes. ONT MinION has been demonstrated to have comparable data quality to the Illumina MiSeq [37]. Low coverage MinION sequencing data can be improved by using a near-identical high quality reference to manually estimate the true sequence with indels being the most common error [37].

3.6 Future outlook for sequencing technologies

Sequencing technologies have mostly been developed to be high throughput with low cost per sample and are frequently located in centralised hubs. There have been significant developments enabling use of sequencing technologies in near patient and/or lower resource settings. There are a number of sequencing platforms under development, with some companies developing platforms designed for use in a rapid setting. However, limited information is available to assess these technologies and their utility for NPPS.

One approach that is of interest for sequencing is the development of solid-state nanopores, which refers to nanopores fabricated from synthetic materials. Unlike biological pores, solid-state nanopores are more stable in terms of diameter and channel length. The biological nanopores used for ONT are limited by their size which allow more bases into the pore. This alters the membrane charge which has an impact on the specificity of the sequencing of these platforms. Solid-state nanopores would be smaller, allowing fewer nucleotides to enter the pore at any one time with the potential to increase sequencing accuracy [23]. ONT are actively researching the use of solid-state nanopores and other technologies to improve their sequencing chemistries.

The sequencing technology landscape has changed rapidly in the last decade with some notable examples of companies no longer providing sequencing platforms. Strong competition and success of certain sequencing companies makes it challenging for new sequencing companies to enter this market. There have also been a number of patent disputes between sequencing technology companies. One notable example is the successful suit brought by Roche's 454 Life Sciences against Thermo Fisher's Ion Torrent related to sample preparation methods prior to sequencing [67]. 454 Life Sciences sequencing platforms are no longer available. A recent verdict against a BGI patent infringement regarding Illumina proprietary azido sequencing-by-synthesis chemistry has restricted the growth of BGI technologies in certain markets [68]. There is interest in how the expiration of certain patents on amido-azyl blocking groups held by Illumina in 2023 will change the sequencing landscape [69].

3.7 Considerations for NPPS

The main constraint related to the sequencing platforms for NPPS are their size, limiting portability, and complex environmental requirements which have proven challenging to achieve within centralised laboratories in LMICs. Limitations around sample transport result in difficulties when returning samples from the field and portable sequencing allows a distribution of this sequencing network [37]. Establishing sequencing technologies requires an 'ecosystem' to be built with equipment that facilitates isolation of nucleic acids to generation of sequencing libraries, as well as the use of algorithms to translate the output of sequencing platforms into sequencing data, assembly and base calling of this sequencing data and further bioinformatics processing. Significant technical expertise is required to facilitate this process and given the rapid evolution and refinement of these technologies, the ongoing use of genomics is dependent upon the development of a sustainable model for this service delivery.

There are a number of considerations for use of NPPS, which will be driven by the chosen sequencing technology and infrastructure requirements. Innovations in sequencing workflows (Chapter 4) and mobile laboratory technologies (Chapter 5) enable the use of sequencing technologies outside of centralised laboratories and in more challenging environments. Regardless, defining the ideal characteristics of sequencing technologies for NPPS enables an understanding of the current gap between technical requirements of available sequencing platforms and the essential or desirable characteristics for NPPS. We will explore this gap further in Chapter 7.

4 Sequencing approaches and workflows

4.1 Sequencing approaches

NGS can be used in a variety of ways to obtain a range of information about pathogen genomes. For example, the aim may be to sequence and identify all microorganisms present in a sample, or focus on a subset of pathogens of interest. The choice of sequencing approach is primarily determined by the needs of the application, the suitability of the method for certain pathogen and sample types, the type of data generated, available infrastructure, expertise and resources, and turn-around time. In theory, any method could be used in a near-patient setting, however there is not a 'one size fits all' solution.

NGS approaches can be broadly divided into two categories based on whether they are pathogen agnostic (i.e. not aimed at sequencing any one specific pathogen) or pathogen targeted (i.e. they are used for sequencing predefined pathogens).

4.1.1 Key parameters to determine NGS approach

The type of genomic information required. NGS can be used to sequence either defined regions or an entire pathogen genome. If sequencing a single gene can identify a pathogen, this may be sufficient for diagnostic purposes. However, for surveillance or outbreak monitoring purposes, a whole genome may more suitable.

Prior knowledge of pathogen. Some NGS approaches require determination of which pathogens and/or genetic targets to investigate in advance of the assay. This could be suitable if the likely cause(s) of a disease is known, for example following an earlier diagnostic test, the presence of well-defined symptoms, or based on knowledge of pathogen prevalence. In other situations, all genomic material present in a sample is sequenced, then analysed to obtain the information required. This can be useful if the pathogen target is not certain, or there is a need to obtain information on an entire microbial community in an unbiased fashion.

Type of pathogen. Features of different pathogens, such as the size and complexity of their genomes (including the presence of mobile genetic elements in bacteria), their amenability to culture, and the presence of homologues genes across phyla, can make them more amenable to some approaches than others.

Type of sample. Some approaches require prior isolation of the pathogen of interest through culture, others can also be used or are only suitable for direct analysis of the clinical sample. Sample types also differ in their suitability for different types of NGS, for example, depending on the pathogen load present and whether they contain high levels of background host DNA (see Section 4.3).

Turnaround time required. Some approaches are comparatively quicker, for example direct-from sample approaches have faster turnaround times than those requiring pathogen culture prior to sequencing. More targeted approaches, which require simpler data analysis, can also help decrease turnaround times.

4.1.2 Pathogen agnostic sequencing approaches

Pathogen agnostic approaches to sequencing are used when the intention is not to sequence and identify predefined microbes; all microbes or a certain class of microbes (such as bacteria or fungi) present in a sample are sequenced. Analysis of the sequencing data is then used to identify, and in some cases obtain WGS information for, any pathogens of interest present in a sample. There are two main ways in which pathogen agnostic sequencing can be achieved, using metagenomic methods which sequence all genomic material in a sample, or targeted NGS methods which sequence common genetic targets

shared by broad categories of microbes. A summary of the advantages and limitations of these approaches is available in Appendix 8.3.

Metagenomic NGS (mNGS)

Genomic material in a sample, which typically contains a mixed microbial community, is extracted directly from the sample and sequenced using unbiased, non-targeted and sequence-independent methods, known as shotgun metagenomics. Whilst mNGS strictly refers the sequencing of all genomic material, in practice often only the RNA or DNA is sequenced, and non-microbial DNA (i.e. human host DNA) is removed. The sequencing data produced is then compared to a database of known microbial sequences to identify the known microbes present and highlight if a novel microbe of interest appears to be present which cannot be matched to the database. Depending on the number of sequence reads produced for each microbe, these may also be assembled to facilitate partial or whole genome sequencing**.**

Example: Use of mNGS to help understand chains of transmission in the Nigeria 2018 Lassa fever outbreak [70]

mNGS was performed upon clinical samples from Lassa virus cases reported at the Irrua Specialist Teaching Hospital (ISTH) between 1 January and 18 March 2018, in order to understand the molecular epidemiology of the unfolding outbreak, understand chains of transmission and identify if a new strain was emerging. 120 samples were selected for sequencing based on PCR analysis of PCR viral load and/or sample origin. mNGS was performed using random reverse-transcription and amplification by Sequence-Independent Single Primer Amplification (SISPA), with libraries sequenced on the ONT MinION. Samples were sequenced at ISTH over seven weeks, with real-time analysis of 36 genomes and subsequent confirmation using all 120 samples sequenced. Phylogenetic analysis was used alongside epidemiological investigations to identify and investigate potential chains of transmission. The authors chose metagenomics as the preferred approach and the ONT platform for NPPS because:

- Metagenomics was better able to capture the diversity of Lassa virus strains than other types of sequencing methods that rely on specific primer design, such as amplicon sequencing
- Metagenomics was able to identify multiple different viruses, allowing the detection of co-infections
- The ONT device used was small and portable so sequencing could be performed on-site, leading to faster results to inform outbreak management.

The analysis revealed extensive diversity and phylogenetic intermingling with strains from previous years, suggesting independent zoonotic transmission events. These results addressed the concern that a new strain of the virus had emerged or that there had been extensive human-to-human transmission, and this information was used to inform local public health control measures.

Advantages and limitations of mNGS

Advantages

- As a pathogen agnostic method, pathogens to be identified do not have to be specified in advance, so sequencing is useful when the likely pathogen identity is unknown, or to identify unusual pathogens present in a sample
- As an untargeted, unbiased method of sequencing, it can be used to sequence and identify all known microbes present in a sample including bacteria, viruses, fungi and protozoa, although limitations may be imposed by sequencing only RNA or DNA.
- It is used for direct analysis of genomic material from samples, without prior isolation of pathogens through culture

- As all genetic material is sequenced, features such as AMR genes are also detected, which can be useful for indicating antimicrobial susceptibility.
- Can be used to provide context on the microbial community and diversity within a sample, including the presence of co-infections
- Capable of fulfilling a variety of purposes, from pathogen identification to producing WGS data, depending on the sample type, amount of microbial DNA present, and the need for sequencing
- Sequencing data can be reanalysed when needed, for example if a new bioinformatics method becomes available, or to carry out retrospective investigations to determine the presence of new pathogens of interest.

Limitations

- As a pathogen agnostic method, metagenomics identifies all microbes present in a sample. When used for diagnostic purposes this could lead to uncertainty when interpreting which results are clinically relevant to an infection, especially in samples with high levels of commensal microorganisms such as respiratory and faecal samples [71]
- As a pathogen agnostic method, can only be used to identify pathogens previously sequenced and where sequence data has been uploaded onto an accessible database. The quality of the database used will impact upon the ability to identify pathogens
- The data produced from metagenomics may be in excess to that required for some of its applications, e.g. for some diagnostic tests it may not be efficient to sequence all genomic material in a sample if the most likely pathogen targets are known, or additional unintended findings may be produced not relevant to the infection diagnosis
- The low prevalence of certain pathogens and gene targets in some samples compared to host DNA can make assay development complex, and limit the specificity and sensitivity of the method for pathogen detection
- Assembling a genome from metagenomics data requires bioinformatics expertise, is computationally expensive and likely to result in lower genome coverage than WGS of microbes isolated through culture or enriched for through pathogen targeted methods of WGS
- Sequencing results can be easily affected by multiple parameters, including sample type, the types of microbes present, sampling strategy to DNA isolation and sequencing methods, making standardisation particularly important, if metagenomics is to be routinely used. This may be challenging to achieve nationally and internationally
- If multiple microbes are detected, it may be hard to assign genetic features such as AMR genes to specific pathogens detected, unless more comprehensive genome assembly can be performed
- As some human DNA is also inevitably sequenced, extra requirements over data handling and privacy may apply.

Pathogen agnostic targeted NGS

Conserved regions of the genome shared by an entire group of microbes are targeted, typically amplified by PCR, and then sequenced. For example, the 16S and 18S ribosomal RNA (rRNA)genes found in bacteria and fungi respectively are commonly used to allow identification of multiple bacterial and fungal species. Analysis of the sequencing data produced is then performed in a similar way to metagenomic NGS, to allow microbes present to be identified based on the sequence data obtained from the targeted genome regions.

Example: Development of 16S rRNA sequencing as a near patient diagnostic test

Laboratory-developed targeted 16S rRNA sequencing tests are already used as tools for bacterial pathogen identification, typically in cases where pathogen identification through culture has failed. Its utility for pathogen identification varies depending on the sample type, with some samples such as cerebrospinal fluid, bone and those from joint infections or abscesses yielding better results compared to culture than other sample types [72]. For example, in the UK 16S testing may be performed in-house in hospital laboratories, but is often outsourced to specialist providers [73, 74]. Therefore, sequencing is not typically performed in a near patient context. However companies such as DNAe are attempting to develop platforms capable of point of care pathogen agnostic targeted sequencing based on 16S rRNA sequencing in bacteria, as well as conserved targets in other microorganisms such as 18S rRNA gene in fungi, simplifying the entire workflow into one automated procedure performed on a standalone device [75]. If innovations such as this are successful, this could help bring 16S rRNA sequencing for diagnostics into more near-patient settings.

Advantages and limitations of pathogen agnostic targeted NGS

Advantages

- As a pathogen agnostic method, pathogens to be identified do not have to be specified in advance, so sequencing is more useful when likely pathogen identity is unknown, or to identify unusual pathogens present in a sample
- Can provide an overview of the microbial diversity present in a sample
- Targeting allows deeper sequencing of the regions of interest, increasing sensitivity for the pathogen targets
- Pathogen agnostic targets can be combined in same assay with pathogen specific targets or other specific targets such as AMR genes
- Suitable for multiplexing, allowing multiple samples to be sequenced simultaneously
- Can provide relatively simple, actionable results on whether a particular pathogen or genome target is present
- Can be suitable for direct analysis of genomic material from samples, without prior need to isolate pathogens by culture. Alternatively, culture could be used to enrich pathogen DNA, but sequencing could be performed on culture broth prior to isolation and identification of pathogens.

Limitations

- As a pathogen agnostic method and similar to metagenomics, multiple microbes present in a sample are identified. When used for diagnostic purposes this could lead to uncertainty when interpreting which results are clinically relevant to an infection, especially in samples with high levels of commensal microorganisms such as respiratory and faecal samples
- As a pathogen agnostic method, can only be used to identify pathogens previously sequenced and where sequence data has been uploaded onto an accessible database. The quality of the database used will impact upon the ability to identify pathogens
- Some groups of pathogens (e.g. viruses) do not typically contain conserved regions so pathogen agnostic targeted methods cannot be used
- Since only small regions of the genome are sequenced, applications of the sequencing data beyond use for pathogen identification are more limited
- Pathogens with limited sequence variability within the conserved genomic regions targeted may not be able to be reliably distinguished from one another.

4.1.3 Pathogen-specific sequencing approaches

Pathogen-specific approaches are those which aim to sequence only pre-defined microbes present in a sample. This can be achieved by targeting and sequencing genomic regions specific to the pathogens of interest, requiring prior knowledge of a target pathogen's genomic sequence. Targeted approaches are typically used in two ways to obtain different extents of genomic information: pathogen-specific targeted NGS methods, which only sequence certain regions of the genome, and pathogen-specific targeted WGS methods, where the aim is to obtain a whole genome sequence. An alternative to targeting the pathogen genome is to culture and isolate the pathogen of interest prior to performing WGS. In this case a non-targeted, sequence-independent method can be used for the WGS of the isolate, as the pathogen has already been selected for via culture.

Pathogen-specific targeted NGS

NGS can be used to sequence genomic target regions belonging to specific pathogens, facilitating their identification and also providing information used for strain typing and antibiotic resistance prediction, depending on the number of targets included for a particular pathogen. Multiple separate targets belonging to different pathogens can be combined in a panel test to allow identification of multiple potential pathogens in a sample. In addition, other genetic targets of interest such as AMR genes can be included in a panel, alongside microbe specific targets. For example, Illumina and IDbyDNA have developed a respiratory pathogen ID/AMR enrichment panel (for research use only) which targets 1,500 markers to identify known respiratory pathogens (187 for bacteria, 53 for fungi and 42 for viruses) and performs concurrent profiling of 1,218 AMR markers. This allows pathogen characterisation from respiratory samples in one assay, with a 24 hour sample to result turnaround time [76].

Example: Seq&Treat project on targeted testing for drug-resistant tuberculosis (TB)

FIND's Seq&Treat project is evaluating the use of different targeted sequencing solutions for diagnosis of drug-resistant TB directly from clinical samples, with a focus on use in LMICs [77]. Targeted sequencing could overcome the limitations of targeted PCR tests which typically can only detect a few drug resistance-associated mutations. The goal is that targeted sequencing will deliver sufficient information to be of similar utility in AMR prediction as the culture-based WGS methods currently used, but with results delivered on a much faster timescale, since culture will not be necessary. TB culture is a slow process that can require 6-9 weeks to obtain drug susceptibility results [78]. The deep sequencing coverage enabled by sequencing targeted genomic regions compared to the entire genome may also increase sensitivity for mutations relevant to drug resistance [79].

The three companies participating in the project to date have developed end to end targeted sequencing solutions [79]:

- GenoScreen's Deeplex Myc-TB test sequenced using Illumina's MiSeq
- Advanced Biological Laboratories' TB diagnostic workflow using Illumina's iSeq
- Oxford Nanopore Technologies' rapid DR-TB detection method using the ONT MinIon sequencer

Each solution is based on targeting multiple genes associated with TB drug resistance, as well as using sequencing of the gene targets to identify the genotype of the TB strain. Whilst these approaches have not been used in near-patient settings and are being evaluated in centralised laboratories, many of their intended features could make them applicable to use in lower resource near-patient contexts in future. These include the culture free nature of the tests, fast turnaround times, user friendly protocols, automated analysis and user-friendly interpretation of the sequencing data, and compatibility with relatively portable and accessible sequencing technologies with varying degrees of throughput [79].

Advantages and limitations of pathogen specific targeted NGS

Advantages

- Targeting allows deeper sequencing of the regions of interest, increasing sensitivity for the pathogen targets
- Targets can be combined to identify a range of different pathogen types from the same sample, as well as key AMR genes present
- Useful for detection of pathogens where pathogen agnostic targeting is not possible due to lack of conserved regions, e.g. many viruses
- Can be used alongside pathogen agnostic targets in the same assay
- Depending on the number of targets used per pathogen, can be used to identify AMR genes and for pathogen typing
- Suitable for multiplexing to allow multiple samples to be sequenced simultaneously
- Can provide relatively simple, actionable results on whether a particular pathogen or genome target is present
- Can be suitable for direct analysis of genomic material from samples, without prior need to isolate pathogens by culture, making it useful for detection of difficult-to culture pathogens, such as viruses.

Limitations

- As a pathogen specific test, the approach is limited to detection of pre-specified pathogens, which limits utility in settings where likely pathogen identity is unknown or uncertain, and pathogens in a sample not covered by the panel will go undetected
- Direct-from-sample sequencing can be challenging, due to limited amounts of pathogen nucleic acids present in some clinical samples
- Evolution of pathogen genomes may mean that targets, for example for diagnostic tests, will require regular updates
- If only small regions of each pathogen genome are sequenced, applications of the sequencing data beyond use for pathogen identification e.g. strain typing may be limited
- If clinical samples contain multiple microbes, it may be hard to assign genetic features such as AMR genes to specific pathogens detected
- If the test is only intended for pathogen identification and the number of pathogens requiring investigation is limited, then targeted PCR tests may be sufficient. Targeted NGS may not be cost-effective if only few targets require sequencing and multiplexing is not required.

Pathogen-specific targeted WGS

If the aim of the sequencing is to perform WGS of a single specified pathogen, an alternative to using culture prior to sequencing is to use a targeted method to allow the enrichment and sequencing of the genome of interest. This means sequencing can be performed directly from the sample, for example through the use of specific PCR amplicons or DNA probes. In this case enough multiple targets are used which cover different regions of the genome of interest, so that when the sequence fragments from each target are pieced together, the whole genome can be assembled.

Example: ARTIC network viral sequencing.

The ARTIC network is a project developing an end-to-end system for processing samples from viral outbreaks to generate real-time epidemiological information that is interpretable and actionable by public health bodies [80]. They have created a lab in a suitcase model to deploy Oxford Nanopore Technologies' MinION sequencing to remote and resource limited locations, and have developed and used several protocols based on tiled amplicon

sequencing for WGS of Ebola virus and SARS-CoV-2; the same method has also been used for sequencing of Zika virus and is applicable to other viral genomes [81].

Advantages and limitations of pathogen specific targeted WGS approaches

Advantages

- As a method for WGS, it provides the most complete pathogen genomic data for a variety of purposes and has the potential to answer multiple questions using a single assay, including genetic typing, identification of AMR genes and prediction of AST, and use of genomic data for surveillance and outbreak tracking purposes
- Targeting allows specific amplification of each genomic region prior to sequencing, increasing overall sensitivity for pathogen detection and ability to produce a whole genome sequence
- As a pathogen-targeted method it is suitable for direct analysis of genomic material from samples, without prior need to isolate pathogens by culture, making it possible to sequence difficult to culture pathogens, such as viruses
- Targets can be combined to identify a range of different pathogen types from the same sample, as well as key AMR genes present.

Limitations

- Evolution of pathogen genomes may mean that target designs will require regular re-evaluation to check they are still relevant and be updated if necessary. Suitability for fast evolving or very genetically variable pathogens, such as the Lassa Fever virus, may therefore be limited
- As targets are required to cover the entire pathogen genome, may be less suitable for pathogens with larger genomes, where it would be complex to design and validate amplicons or baits to cover the entire genome. For bacteria, culture remains a simpler method to increase the amount of genetic material available for WGS
- More complex genomic data is produced and genome assembly requires a greater understanding of bioinformatics and quality control procedures
- Use of data for purposes such as detecting chains of transmission in outbreaks or tracking pathogen evolution may require further specialist expertise and training.

Unbiased WGS of culture isolates

The pre-sequencing step of culture is first used to isolate a single pathogen and typically to identify it, with this step also acting as an enrichment step for the target pathogen's DNA. DNA extracted from the culture isolate is then sequenced using unbiased methods as in metagenomic sequencing, however in this case it is known that only one pathogen will be identified from the sample. Although the sequencing method is unbiased and not targeted to a particular pathogen, the overall sequencing approach is one carried out with the aim of further characterising a specific known pathogen.

Example: Implementation of WGS in a diagnostic laboratory in Australia [82].

In 2017, a hospital diagnostic laboratory in New South Wales, Australia, instituted a pathogen WGS service to sequence bacterial isolates within the hospital's diagnostic laboratory. The sequencing service serves the 900-bed referral hospital. The hospital laboratory currently performs pathogen WGS on approximately 1000 bacterial isolates per year, primarily using an Illumina MiSeq. WGS is performed on culture isolates, which have been flagged for sequencing following identification by MALDI-TOF, susceptibility testing, and PCR confirmation. WGS has been used for individual patient decision making, hospital decision making, and investigating unexpected diagnostic test results [82]. The hospital found that the main advantages of in-house sequencing over sending their samples to an external centralised laboratory for sequencing were:

- Reduced turnaround times as samples did not have to be transported to external sites, which was especially useful for outbreak analysis
- The ability to prioritise which samples needed more urgent sequencing
- A better ability to analyse genetic data from samples in relation to their associated metadata and to other isolates of the same bacterial species previously sequenced, enabling connections between samples to be established which may not have been apparent when samples are individually sequenced at a centralised laboratory
- Better communication between clinicians, microbiologists and bioinformaticians allowing better interpretation of results
- The ability to sequence single isolates which could help early identification of an extended outbreak; typically a cluster of likely linked isolates is needed to warrant sequencing at a centralised facility.

Advantages and limitations of unbiased WGS of culture isolates

Advantages

- As a method for WGS, it provides the most complete pathogen genomic data for a variety of purposes and has the potential to answer multiple questions using a single assay, including genetic typing, identification of AMR genes and prediction of AST, and use of genomic data for surveillance and outbreak tracking purposes
- Culture is a relatively simple and established way of enriching pathogen genomic material for WGS (where facilities are available), enabling more high quality and comprehensive sequencing data to be produced
- Unbiased sequencing of culture isolates can be used to identify unknown pathogens or provide sequence data of novel pathogens, if it has been possible to culture the pathogen when it could not be identified via other methods (such as gram staining, microscopy, biochemical tests or mass spectroscopy)
- Comprehensive data enables re-analysis for the presence of new genomic features of interest that become applicable e.g. to retrospectively track the emergence of a novel genetic variant.

Limitations

- Culture requires additional facilities and expertise, and not all pathogens are amenable to culture
- Culture can select the pathogen/strain most adapted to growth in the culture media, not the most virulent or pathogenic
- During the culture process some genetic characteristics of the original pathogen population infecting the patient may be lost
- The time taken to obtain a culture isolate increases the overall turnaround time of producing a genomic sequence from a sample
- More complex genomic data is produced and genome assembly requires a greater understanding of bioinformatics and quality control procedures
- Further analysis and interpretation of data for purposes such as detecting chains of transmission in outbreaks or tracking pathogen evolution will often require further specialist expertise and training.

4.1.4 Conclusions

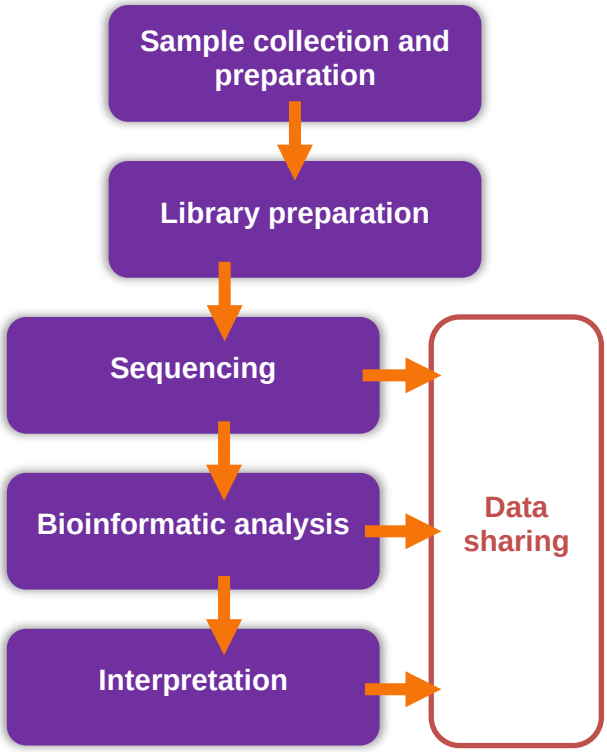
A range of approaches can be used for NPPS, suitable for producing different types and amounts of genomic information. A key choice is whether a pathogen-agnostic or pathogen-specific approach is most suitable, and for the approach taken what extent of genomic information is required. The main advantages and limitations of the pathogen-agnostic and pathogen specific-approaches described above are summarised in Appendix 8.3. Factors that influence the choice of technical approach for NPPS scenarios include:

- The ability to perform culture-free sequencing in settings where culture facilities are unavailable or where there are a diverse range of pathogens requiring sequencing including those not amenable to culture
- The level of expertise needed to perform the sequencing and interpret the results, in some near-patient settings such as diagnostic laboratories there will likely be no specific expertise surrounding sequencing and relatively simple and user-friendly approaches will be needed
- The turnaround time, with culture-free methods in particular offering faster turnaround times than those requiring culture, which may be especially important in diagnostic settings or outbreak scenarios.

4.2 Sequencing workflows

Sequencing workflows are highly flexible with different kits and technologies available that allow the user to tailor the workflow to their needs. The sequencing workflow can be built as a single unit or as a series of components. There are a number of choices at each stage of the workflow including decisions determined by the sequencing approach. The complexity of these steps varies depending upon a number of factors, including: sequencing equipment selected, sample type and purpose of sequencing.

Figure 1. Key steps in a sequencing workflow



Key steps in a basic workflow can be summarised as follows:

- **Sample collection and preparation.** Includes the steps from collection of a patient sample through to the storage and transportation of that sample, to the extraction of nucleic acid prior to further processing. It can also extend through to conversion of RNA to cDNA, depending upon the workflow and the pathogen being sequenced

- **Library preparation.** Transforms the retained nucleic acid portion of a collected sample into a prepared sample library ready for sequencing. This may include fragmenting or size selection of nucleic acids (dependent upon application), addition of sequencing adapters and quantification and quality control of resulting libraries
- **Sequencing.** The process by which the sequence of bases in a series of nucleic acids is detected by one of a number of methods to provide readable data – raw sequence reads. This process covers entry of a prepared sample library into a sequencing system to retrieval of raw sequence information
- **Bioinformatic analysis.** This includes the processing and conversion of raw data that is produced during sequencing into one of several formats that is suitable for ongoing analysis or interpretation.

Data sharing may also be considered part of this workflow. The approaches adopted towards data sharing, including the databases used, file type, and attachment of metadata will impact upon the ability of others, outside of the original sequencing group, to effectively use the information.

4.3 Sample consideration and sample type

Depending on the pathogen of interest, a range of sample types can be collected for sequencing using a variety of methods. These methods will differ according to ease of collection, storage, and sample degradation over time. Certain samples will be more appropriate in low resource settings, for example external swabs or blood finger-prick samples, whereas safely retrieving some other types of samples (e.g. cerebrospinal fluid) can be more difficult.

Methods used to store samples may cause nucleic acid degradation, for example, freeze-thaw cycles due to an unreliable cold chain, or extremes of temperature or humidity. Direct sequencing of samples can help reduce the likelihood of sample degradation prior to sequencing. Storage mediums are available that protect DNA and RNA in samples (e.g. RNA Shield).

Any sample can be used for sequencing, but certain samples may be more or less difficult to sequence due to high concentrations of host DNA, high microbiome background or presence of contaminants. Choice of sample type will be driven by considerations related to the quantity of target nucleic acid within the sample. This will affect the DNA extraction method and choice of sequencing approach. In a clinical setting, direct sequencing is desirable but may require additional stages including, for example, host depletion, target enrichment, and cDNA conversion. As a result, sample type and sequencing approach decisions will need to be aligned to the purpose of sequencing. These decisions will affect the analytical sensitivity and specificity of the sequencing workflow [83].

Culture methods are used to increase the amount of sample DNA and isolate the pathogen of interest. However, these methods will increase the turnaround time of the sequencing workflow. Sequencing from isolates may be advantageous where sequencing is being performed alongside traditional microbiology methods, such as antimicrobial susceptibility testing (AST).

Any infectious disease sample will need to be handled carefully prior to inactivation with further sample processing under biosafety level (BSL)-2 conditions [36]. Considerations for safety include personal protective equipment (PPE), disinfection and inactivation protocols, and waste management protocols.

4.3.1 Biosafety laboratory and sample transport regulations

Laboratory biosafety measures should be put in place following the appropriate risk assessments to minimise the likelihood and severity of any potential exposure to pathogens and ensure a safe workplace [84]. There are four protection levels of biosafety to handle

pathogens and the nature of the pathogen will define the level of biosafety required. If the pathogen is infectious to humans, the mode of transmission and severity of disease if infected will inform the measures needed to protect those handling pathogen samples [41].

Samples that are transported between the patient and the laboratory need to be transported in accordance with UN3373 requirements, regulations relating to packaging requirements for biological and infectious substances [85]. UN3373 Category A present a high risk to individuals and community and are defined as an infectious substance capable of causing permanent disability, life-threatening or fatal disease to humans or animals. UN3373 Category B are specimens that do not fall into category A.

4.4 Sample preparation

Sample preparation involves processing a sample and preparing its nucleic acid for sequencing. Key stages include [69]:

Nucleic acid extraction. The method used will vary depending upon the sample type, the quantity of nucleic acid and whether the target nucleic acid is DNA or RNA. Different extraction methods will vary in terms of ease and speed of use, and cost per sample. Choice of method will require consideration of these factors, and the context of use compared to the volume and purity of nucleic acid required to undertake the analysis.

Purification and quality control. Once the DNA or RNA has been extracted, cellular debris and other contaminating material will need to be removed. Different methods are available for these stages, typically done by magnetic bead-based clean up or an agarose gel. Host depletion methods can be used to remove human DNA. An important quality control step is determining nucleic acid concentration and is the final step before sequencing. For NPPS, spectrophotometric, electrophoresis and PCR based methods can be used to determine the quality and quantity of nucleic acid [86]. Additional amplification stages may introduce artefacts and unnecessary amplification should be avoided where alternative methods are possible.

Library Preparation. Library preparation converts the nucleic acid into an appropriate format depending upon the chosen sequencing technology. This process includes a number of stages. For short-read sequencing technologies, fragmentation is required. For all sequencing, attachment of adaptors and library quantification is required. Additional stages will include: barcoding samples, targeted amplification or enrichment (depending upon the sequencing approach discussed in above) and complementary DNA conversion of RNA for some sequencing technologies. Different library preparation kits are available from a range of commercial suppliers, with different advantages and limitatioms depending upon the sequencing approach, choice of sequencing technology, speed of method, cost of kit, ease of use or hands-on time and use of automation:

- **Targeted sequencing**: Broad category of techniques where target enrichment is used to extract specific regions of DNA or RNA for sequencing. This method will generate a smaller dataset typically reducing the complexity of analysis. Broadly, methods fall into categories: hybrid capture based target enrichment, amplicon-based target enrichment and molecular inversion probes (MIP)
- **RNA sequencing**: Currently a necessary approach to sequence some viral genomes on certain sequencing platforms (e.g. ONT MinION). Specific attention needs to be taken at sample collection when RNA is the focus since the nucleic acid can quickly degrade. Additional use could be for more complex sample types where RNA may be more abundant than DNA where the pathogen is transcriptionally activated suggesting a greater role in the patient's illness. RNA sequencing may be one strategy to enrich for the pathogen.

4.4.1 Considerations for sample preparation

A key consideration in establishing NPPS workflows has been to simplify them and reduce the number of steps where possible to reduce any risk of errors. Current sequencing platforms have a range of sample preparation kits available which may differ in terms of required processes, number of steps and the resultant turn-around-time. Automation will reduce human error in the sequencing workflow. This is an area of ongoing research and further development is needed. To simplify workflows for NPPS, reagents have been developed that do not require storage in a fridge or freezer (e.g. ONT Field Sequencing Kit), and some processes now have lower equipment requirements, for example methods for DNA extraction without a centrifuge [87]. Expert interviewees indicated that they are largely using rapid workflow kits in near patient settings, however, sequencing will be performed using any available kits and protocols within a group or country.

Ideally, some optimisation should be performed to evaluate reagent performance and ensure the kit meets the workflow requirements. Nucleic acid extraction protocols will have different performance for different sample types. Additionally, these protocols may be better for DNA or RNA protocols and, therefore, kits used within a workflow should be selected depending upon the intended use. Innovations around laboratory equipment have miniaturised these devices, including centrifuges and PCR machines, small enough for use in, for example, a suitcase laboratory [88].

4.5 Amplification and enrichment methods

4.5.1 Polymerase chain reaction (PCR)

Nucleic acid amplification is a critical component of many sequencing workflows, needed to ensure sufficient genomic material when sequencing. Traditionally, amplification has been performed using polymerase chain reaction (PCR), a process that performs exponential amplification from a single nucleic acid through a series of cyclical temperature changes that support the different chemical reactions involved in the process: strand denaturation, primer annealing, and enzymatic extension [89]. PCR is therefore dependent upon the presence of a thermocycler to provide these successive heating and cooling cycles. Small PCR machines are available – for example, the miniPCR TM mini8 thermocycler is an 8-well thermal cycler and has been used successfully in remote locations to perform PCR [90]. Another example is miniPCR, which can be used to prepare the libraries and amplicons and there is also the option of the MyGo mini S real time PCR instrument to quantify the DNA (16 samples at a time).Innovations have reduced the size of PCR machines making them more amenable to use in remote settings with demonstrable utility in diagnostics for RT-qPCR and as part of portable sequencing workflows [91].

4.5.2 Isothermal loop amplification methods

Alternative technologies to PCR are being developed to enable amplification of nucleic acid without the need for thermocycling and removing the need for this equipment. These methods are designed to be targeted, with specific primers developed to the target sequence of interest. This is based on isothermal amplification which is a method that allows exponential amplification of nucleic acid at a constant temperature [92]. A number of different methods have been developed, including loop mediated isothermal amplification (LAMP), nucleic acid sequence-based amplification (NASBA) and recombinase polymerase amplification (RPA). These methods broadly use different enzymes or nucleic acid primers to separate DNA strands for synthesis of complementary strands, removing the need for the high temperatures required for this step during PCR-based methods. Isothermal amplification may be faster than traditional PCR approaches because this process is not linear, dictated by the heating and cooling cycles; instead these reactions can occur simultaneously resulting in more rapid amplification of the target sequence [89]. These

reactions work best within a defined temperature range and may require a water bath or other technique to maintain the optimum temperature.

Isothermal amplification has also been applied successfully for human whole genome sequencing, single cell sequencing and ancient DNA sequencing [93]. The SARS-CoV-2 pandemic has resulted in significant investment into isothermal amplification methods as a diagnostic with demonstrable success, such as the LamPORE assay from ONT [94]. Field sequencing studies have successfully used isothermal amplification within a suitcase laboratory designed for sequencing workflows [60, 88]. Research is needed to demonstrate if isothermal amplification is comparable to PCR in terms of specificity and cost-effectiveness. These methods have been shown to be insufficiently specific for some primer binding resulting in erroneous product amplification with high background noise. Further research is needed to refine them. Additionally, many methods are protected by patents or licensing by a single proprietary company limiting access to enzymes and reagents underpinning this technology.

4.5.3 ONT adaptive sampling

Adaptive sampling is a computational enrichment method unique to nanopore sequencing platforms. This method allows the user to select target regions of interest, perform real-time sequence to select 'on-target' sequences and eject 'off-target' molecules. This method has been applied to sequence low abundance microbial species in a metagenomic sample [95]. The key advantage of this method is that it is PCR-free, reducing bias that may be introduced when using amplification-based enrichment methods. This method has been applied successfully in human genomics (such as virtual cancer gene panels), where selection and identification of the regions of interest is more easily mapped to a reference genome. Further refinement of this method is needed for sequencing microbial samples to account for complex genomic architecture (i.e. mobile genetic elements), which will support identification of key genetic features in the sample of interest. For NPPS, this method has the potential to reduce the volume of sequencing required to obtain sequence data for a specific purpose.

4.6 Data analysis

Regardless of the location where sequencing is being performed, the data generated needs to be processed through a number of stages to enable analysis and interpretation. These stages of data processing include quality control, assembly, and sequence annotation. A number of quality control stages need to be built into the bioinformatics pipeline to ensure the accuracy and precision of any downstream analysis [96]. These steps include evaluating the raw sequence data, trimming reads, identifying contamination and setting quality control (QC) parameters [96]. Depending upon the sequencing approach, different data cleaning stages may be required. For example, host DNA is a major source of contamination from metagenomics and bioinformatics is used to remove these data from downstream analysis.

Bioinformatics pipelines are designed for a particular purpose with different software or tools available to perform these types of analyses. There are a number of possible types of analysis or interpretation including species identification and subtyping, identification of resistance genes or variants, and the study of transmission dynamics. Different tools are available for each of these applications, depending on the pathogen of interest. Many of these tools will require some knowledge of programming in order to write code capable of running automated data manipulation and analyses. There are a number of choices when developing a bioinformatics pipeline, and as a result bioinformatics approaches remain highly variable [97]. Therefore, the choice and customisation of these analysis pipeline(s) will have a profound effect on the interpretation of genomic information.

Establishing a bioinformatics pipeline for NPPS in lower resource settings presents specific and complex issues. Cloud-based solutions, such as EPI2ME from ONT and Illumina

Analytics Connected, enable users to manage data centrally and organise it in a secure environment. Users are then able to run ready-to-use pipelines provided by these tools or develop custom pipelines. This approach is advantageous for data sharing and collaboration between public health authorities or research institutes as required. However, internet connectivity may be unreliable or non-existent and work is ongoing to develop reliable pipelines that can be used off-line. Many tools are accessible primarily through cloud platforms and slow internet connection will increase the turn-around-time of these protocols [43].

The biggest barrier relates to bioinformatics expertise. Expert interviewees highlighted the impact of commercial products to implement a reliable bioinformatics pipeline removing the need for this expertise within the team. In the future, improvements to the web-interfacing and establishment of bioinformatics protocols for NPPS will enable the use of sequencing by non-experts within lower resource settings. Out-of-the-box solutions will improve accessibility of sequencing to address specific clinical questions, including for diagnostics.

4.6.1 Databases and interpretation

In near patient settings, sequencing may be performed in isolation or alongside other tests, for example microbiology or PCR. The ability to interpret the data and the confidence in the result will be limited by the sequencing approach and the quality of the sequence data. More targeted approaches capture specific genetic elements and will limit analysis to these regions only. By comparison, metagenomics captures all genetic information within a sample, however, analysis will be limited to the contents of available databases. Databases are typically developed to address a specific purpose and will often be focused on a particular disease or disease area – for example, salmonella, tuberculosis, malaria or antimicrobial resistance genes. These databases require ongoing curation and maintenance to ensure sufficient quality and relevance to enable use for comparative analysis. There may be more than one database available for a given pathogen. This results in the need to combine analysis approaches to make full use of databases, increasing the complexity of the analytical process.

Diagnostics using sequencing is complicated by the high sensitivity of this method resulting in the potential to identify more than one pathogen, resulting in difficulty distinguishing the disease causing pathogen(s) in this instance. Ambiguous phenotypes, presence of opportunistic infection and poor understanding of these correlations make it challenging to develop diagnostic algorithms using sequencing data.

4.7 Process control

NGS can be affected by a range of artefacts that arise during the library preparation and sequencing processes, which can negatively impact the quality of the raw data for downstream analyses [98]. Sequencing consists of a ‘wet-bench’ and ‘dry-bench’ integrated workflow that is challenging to assess using standard quality metrics used by microbiology laboratories [99]. Standardisation, validation, and quality management of sequencing is a requirement to ensure appropriate process control and confidence in sequence data for any ongoing analysis. It is important to consider process control and the steps necessary to ensure the quality of sequencing data regardless of the location where sequencing is being performed.

4.7.1 Laboratory test validation

Laboratory test method validation refers to the process of evaluating the performance of a new instrument or test methodology in relation to an instrument or method in current use [100]. Principles and guidelines for test validation, quality control, proficiency testing and reference material are available for public health microbiology laboratories [99]. Laboratory test validation is key to ensuring the standards and quality of available tests. For NPPS, the

validation process should compare an established method, e.g. microbiology methods for pathogen identification, against the new method. Experts highlighted that this is a model that is being considered to enable in-the-field validation of sequencing findings. This testing process also enables identification of challenges associated with in-the-field sequencing, refinement of methods and quantification of the interpretation limitations.

Laboratory test validation and development of sequencing workflows is a complex process requiring a high level of expertise and resources. Experience of sequencing workflow validation may be more variable in lower resource settings. However, this is an essential process to test different aspects of the workflow. In addition, performing sequencing using this validated workflow in a central laboratory and a field setting enables comparison of the data to quantify the impact on sequencing data quality and further determine the likely impact on analysis and interpretation. This could involve comparison with sequencing data from a central laboratory. Once a workflow has been validated, programmes can be established to train laboratory personnel using these workflows, establish proficiency testing to assess competence and implement quality assurance programmes as part of ongoing monitoring of sequencing.

Quality assurance and quality control. These programmes are implemented to ensure high quality and consistency of routine sequencing performed in the laboratory and guarantee the process to generate data of sufficient quality for analysis and interpretation [99]. These processes have mostly been established in reference or large microbiology laboratories, and the exact nature and requirements for these programmes will vary depending upon the expertise, regulation and management of microbiology laboratories. Principles of quality assurance and quality control underpin the reliability of sequencing regardless of the location of use (for example hospital, mobile or field laboratory) and the infrastructure of the laboratory (e.g. lab in a suitcase, mobile sequencing vehicle or fixed laboratory). Training to ensure staff understand the importance of these quality assurance programmes will ensure buy-in and sustained integration ultimately ensuring the quality of NPPS. This is most important where sequencing is being performed to support clinical decision making.

Proficiency testing. Also known as external quality assessment (EQA), this process involves the introduction of samples of known but undisclosed content into a laboratory's routine testing procedure [101]. This allows independent assessment of test performance to ensure quality, harmonisation, comparability, and reproducibility of diagnostic results [97]. For NPPS, proficiency testing will ensure that the laboratory has been set up correctly, the proficiency of the staff performing sequencing and the validity of any results. This is critical to ensuring confidence in results and their use in clinical decision making.

Ultimately, process control underpins the utility of NPPS and will need to be considered when developing sequencing workflows for use regardless of the exact setting. Investment to provide reliable workflows with robust process control and quality assurance will create resilience in the system for sequencing to be deployed for NPSS.

4.7.2 Sequencing laboratory workflow

Different sample types and extraction methods can introduce inhibitors or contaminants to the enzymatic reactions for sample preparation and sequencing. Sources of nucleic acid outside of the sample of interest may affect the sequencing run and data quality. To manage this, the workflow requires distinct workstations within the laboratory or equivalent to minimise the risk of contamination. This often requires up to three distinct workstations that are clearly designated for specific tasks, namely:

1. Sample preparation and extraction
2. Mastermix preparation
3. Sequencing preparation – where the mastermix and samples are combined for sequencing

Considerations for the practical implementation of sequencing workflow zoning will be considered in Chapter 5. The sequencing workflow is an integral and vital component of NPPS. All aspects of the workflow will need to be considered and tested as part of implementation in each specific context in which NPPS is being delivered.

5 Mobile laboratories

Mobile laboratories are laboratories that are either fully housed within or transported by a vehicle or person to be set up and used in a non-standard (semi-permanent or temporary) laboratory structure. They can provide rapid response laboratory services in situations such as emergencies due to natural or other disasters, disease outbreaks, or in support of routine surveillance needs. While not all mobile laboratories carry out near patient testing, and not all near patient laboratories are mobile, they nevertheless provide useful exemplars for exploring how to optimise laboratory and sequencing capabilities in resource limited and near patient settings.

Reference or centralised laboratories can be distant from the sites where outbreaks occur, resulting in long sample transport times and delays in results and subsequent management. The solution to the time delay for testing potentially highly contagious infectious pathogens is to have the capacity for this testing at the site of the outbreak. While having a fully equipped biosafety level 3 or 4 (BSL-3 or 4) laboratory in many locations is unrealistic, bringing a mobile laboratory to the area of an outbreak to provide these services is feasible. As such, mobile laboratories have been used for the surveillance and detection of infectious disease as part of research programmes and outbreak responses for decades and over this time technological advances mean that these facilities can carry out many of the procedures currently available in centralised laboratories [102].

During the 2013-16 Ebola outbreak in West Africa, the lack of laboratory capacity within many of the affected countries in the early stages of the epidemic was considered to have contributed to the rapid spread of the disease [103]. Several countries were involved in the deployment of 27 mobile laboratories to provide rapid-in country diagnostic testing, as well as genomic epidemiology, which supported healthcare systems and the public health response [103]. The experiences from the Ebola epidemic resulted in the the establishment of an integrated network of support laboratories providing epidemic preparedness and response capabilities to future outbreaks for infectious diseases in these countries [103].

5.1 A modular approach to mobile laboratories

Mobile laboratories are developing in such a way that many stakeholders and users view them as being modular, where ‘modules’ carry out particular functions that can be deployed and used as needed for a specific situation. The use of laboratory modules supports flexible laboratory design and can also enable targeted responses by including specific diagnostic, clinical or surveillance procedures needed for a particular situation, as well as allowing one laboratory to focus on more than one pathogen or purpose.

To address inconsistencies in the standardisation and functions of mobile laboratories, guidelines on minimum standards and coordination of efforts are being developed. The WHO European regional office recently (2021) developed guidance on rapid response mobile laboratories (RRML) classification [104]. This guidance provides information on key components required to provide standardised mobile laboratory services that can be integrated into wider rapid response activities. This was a process initiated as part of the Global Outbreak Alert and Response Network (GOARN), a WHO network of over 250 technical institutions and networks globally that respond to acute public health events with the deployment of staff and resources to affected countries [105].

The RRML guidance aims to adress a number of the complexities involved in a system such as mobile laboratories [104]. This guidance takes into account different quality management systems and minimum quality standards, as well as the required laboratory information management systems (LIMS) and processes for exchange of data. It considers what variables discriminate the various types of RRMLs including their capability and throughput, biosafety and biosecurity considerations, means of transport, and logistical and operational support requirements (Figure 2). It also contains considerations around well defined

diagnostic modules that can be added or incorporated into mobile laboratories to ensure flexibility, but also highlights measures to ensure that these modules are interoperable and scalable to the required response (Figure 3). Advanced modules include procedures that are more resource intensive in terms of laboratory architecture, consumables and working time, for example high-throughput sequencing (HTS), long-term sample storage and pathogen culture.

Figure 2. The WHO European regional office classification of rapid response of mobile laboratories (RRML) has identified five types of RRML and the variables that discriminate between them [104].

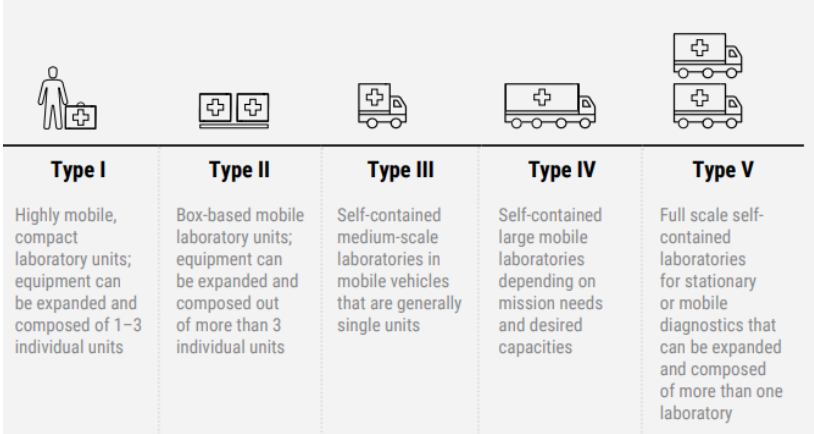
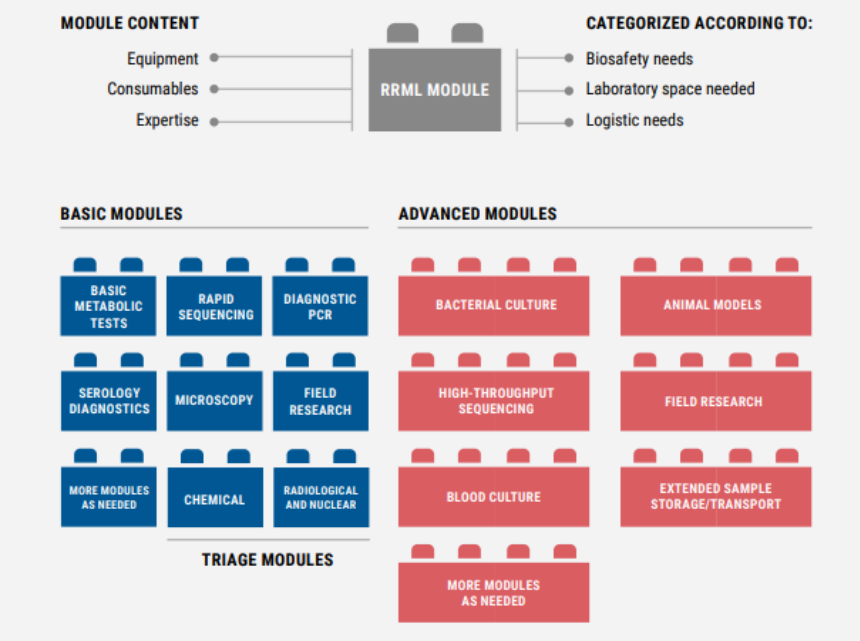


Figure 3. Example modules for configuring RRMLs for deployment from the WHO European regional office classification system [104].



5.2 Key considerations for mobile laboratories

There are a number of infrastructural and logistical challenges associated with the setup and ongoing operation of mobile laboratories [4, 103]. Managing these requires planning and development pre-deployment and ongoing management during day to day operations once

deployed. The development stage requires a multidisciplinary team that could include scientists, engineers, clinicians, health and safety officials, and bioinformaticians. Once out in the field there will be limited opportunities to obtain additional equipment, supplies or consumables so careful planning is required.

Factors to be considered when establishing and operating mobile laboratories include:

Staff support and personnel requirements. When identifying personnel for deployment, particularly in emergency situations, their training and expertise as well as their ability to adapt to a range of potentially challenging circumstances – including potential social unrest – whilst in the field should be considered [102]. During an outbreak, first responders can often work long hours during their deployment, which could extend to weeks, if not months, in duration. Suitable working conditions for staff that include measures against extremes in humidity or temperature as well as associated stressful circumstances that can affect physical, mental well-being and attention to detail are vital. This can also include easy to use assays and automation which reduce psychological load. Other measures to support wellbeing include access to the site, accommodation, food, water and medical facilities, if required.

Training packages. Training for staff is required on a number of levels, to operate and establish the laboratory and ensure ongoing smooth operation. Further training modules, as needed during deployment, can also provide support for sustainable, uninterupted operations. Technologies and protocols that are easy to implement and train can enable this. Training of non-laboratory staff should also be considered, such as technical, operations or logistical support training. One example of this is training the drivers of vehicles to carry out maintenance on the vehicles and some of the laboratory equipment.

Laboratory guidelines and standards. All tests conducted in mobile laboratories should be done under Good Clinical Laboratory Practice (GCLP) guidelines and managed using data management protocols. It should be common practice for mobile laboratories to meet predefined standards and adhere to the same quality assurance mechanims that apply to standard stationary laboratories, ensuring compliance with both national and international standards [104]. However, there are also unique considerations that may be required for mobile laboratories as well as standard operating procedures that require development. For example, guidelines on the calibration of machinery after its been transported to a site, or how to determine levels of vibrations when the laboratory is being transported and the effect this has on equipment both when it is on standby but also when it is in use.

Equipment. The suitability and field usability of equipment requires careful assessment. Equipment needs to function accurately and reliably – according to manufacturer's guidelines – under challenging conditions, which will include variations in temperature, vibration, humidity and dust levels. Most equipment will require field-testing and evaluation prior to use. Mobile laboratories also have space limitations so the footprint and height of an instrument can impact on whether it is included or not. Standard equipment available in static laboratories also needs to be considered, for example hand and eye washing stations, power supply and water purification systems. Alternatives to larger equipment items may be available, for example 'glove boxes', a sealed protectively lined compartment containing access holes with attached gloves allowing handling of dangerous materials inside the compartment, to replace fume hoods [1, 106, 107]. Alternatives include tented glove boxes [108, 109]. Other equipment considerations include personal protective equipment (PPE) for laboratory staff and laboratory consumables, including storage requirements (e.g. refrigeration) for these.

Transport of the laboratory to the site. Access to the affected region is a potential obstacle. While many mobile laboratories have been designed to fit into a varying number of durable crates, others utilise multiple large vehicle containers to provide the necessary infrastructure and equipment, when there are roads that can accommodate them. It is not only transportation of the laboratory but the consumables and reagents as well as staff and

their needs (food and water) that must be considered. Road access leading to the affected area is not always possible and could only be suitable for a motor bike that can transport a technician carrying a laboratory in a backpack. This reinforces the need to develop assays that are independent of traditional laboratory equipment and that can be carried by an individual [102].

Infrastructure status of the mobile laboratory site. Some equipment may require a standing structure and a reliable power supply. There are a number of ways that electrical power can be delivered e.g. via generators, car batteries or solar panels. Mobile laboratories can be set up in pre-existing structures that have all the necessary amenities, for examples schools or unused rooms in clinics. However, it is not uncommon for the infrastructure availability to be unknown and only determined on arrival. This means there may be a requirement for the deployment of tents or containers to house the laboratory [102].

Laboratory design. A design that is adaptable is a benefit. There will however be standard requirements such as the workspace needs, workflow through the laboratory, and physical layout. These should all also consider sample flow, especially receiving of potentially infectious samples, provision of clean and dirty laboratory areas, waste management and measures to prevent contamination.

Sample logistics. Logistical issues in terms of local sample collection and transportation to the laboratory should be considered even in near patient and/or field settings. Appropriate sample documentation and tracking processes should also be in place.

Waste management. A waste management strategy will be required for all laboratory waste, including used PPE. Options include disinfection and disposal of waste on site, or a strategy for safe storage and removal to an external site for disposal.

Data management. Strategies for managing data are particularly important in areas where internet access and/or the mobile phone network access may be unreliable or non-existent. This may require additional equipment for connecting to satellite networks and cloud services, or off-line data analysis protocols and measures to upload data at the end of a deployment.

Societal and cultural considerations. Mobile laboratories, and the clinical staff that work with them, may face challenges that include social or cultural misunderstanding, issues around mistrust and misinformation, or social unrest. This could affect the day to day work of the laboratory, including sample collection, and affect the safety and security of its personnel such that access to security personnel may be required. This adds another layer of complexity to the running of a mobile laboratory during emergency or outbreak responses and as such, the personnel must be physically and mentally fit when deployed [102].

Capacity and sustainability. During the Ebola outbreak many of the laboratories that were implemented were operated by teams of volunteers deployed on a rotation of 4-6 weeks from research and public health laboratories around the world. A common pattern in these situations is equipment being left unused after the outbreak, or removed from the country because local staff lack the necessary training, sustainable funding; affordable reagents and equipment are also not available [103]. To overcome this research groups with mobile laboratory capabilities should establish research projects with local researchers or surveillance programmes with public health officials. This would help nations maintain and expand skill sets to develop local capacity in regions where outbreaks are likely to occur [102]. Ensuring there are training programmes prepared for when a mobile laboratory is deployed could facilitate ongoing local capacity to continue running the laboratory. Capacity building can occur during non-emergency/outbreak periods with the repurposing of rapid response laboratories:

- Laboratory testing could be redefined and mobile laboratories that were used during an epidemic transition from acute testing for clinical triage to routine surveillance testing. Basic laboratory set-up provided for outbreak response could be extended to

develop essential clinical and public health laboratory services for other infectious diseases. Minor modifications of protocols and procedures would make it possible to establish diagnostic tests for a number of endemically circulating pathogens of clinical significance.

- Many of the laboratory techniques used to diagnose a patient can be applied to environmental and/or animal sampling and could be used for surveillance of endemic pathogens contributing to One Health implementation and surveillance.

Having these operating during non-emergency periods would ensure there are resources and a pool of skilled professionals that could be utilised during emergency response periods.

5.3 Mobile laboratory uses

Numerous lessons have been learned from mobile laboratories that have been operational, many of which are now considering how to integrate sequencing; below we outline some examples:

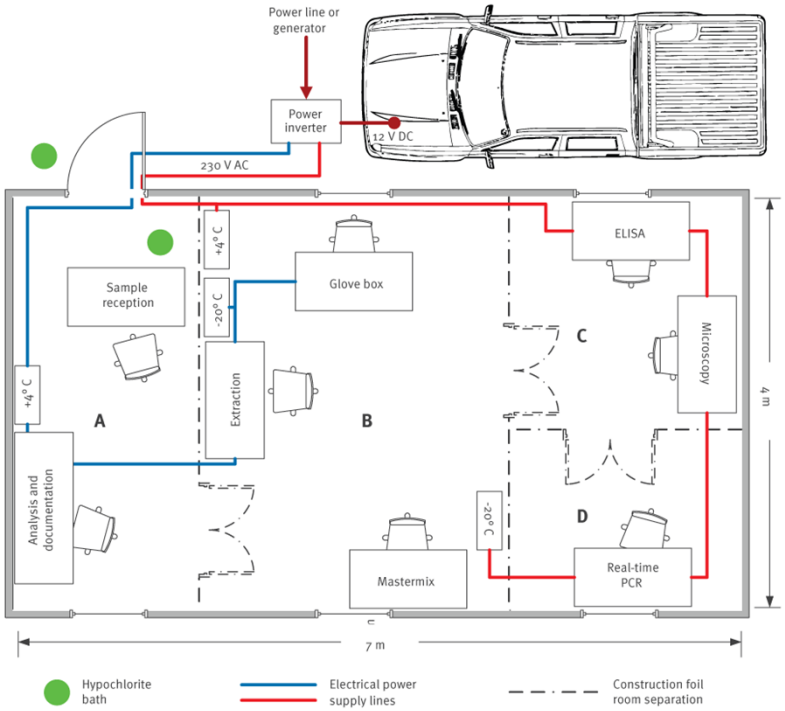
5.3.1 European Mobile Lab during the 2013-16 Ebola epidemic

The European Mobile Lab (EMLab) Project provides a modular boxed system that has been adapted for outbreak response missions in Africa and Europe. It is portable – much of the equipment can be carried by hand – and there are three laboratory units that have been set up by the partners of the Bundeswehr Institute for Microbiology in Munich. One laboratory unit was established at the Institute for Lassa Fever Research and Control at the Irrua Specialist Teaching Hospital, Irrua, Nigeria. The second unit was established at the National Institute for Medical Research in Dar es Salaam, Tanzania, by the consortium partners of the National Institute for Infectious Diseases, Istituto Nazionale Malattie Infettive L. Spallanzani. The third unit was established in Europe at the Bundeswehr Institute of Microbiology. This unit is staffed by the European partners of the EMLab consortium and was ready to respond to outbreaks that could not be reached by the two units in Africa. All three laboratory units were deployed to the 2013-16 Ebola virus disease outbreak in West Africa.

To overcome the problem of the long distances to the central diagnostic laboratories during the Ebola epidemic in West-Africa, the EMLab established a moving laboratory unit in Nigeria. This consisted of 27 boxes, each approximately 20 to 30 kilograms, containing more than 400 equipment items needed to set up a fully-functional BSL3 or BSL4 diagnostic laboratory in a tent or in a local house. Minimum requirements were at least 28 square meters of space and a car for constant energy supply via a power inverter (Figure 4). EMLabs enabled sample inactivation and preservation, molecular diagnostics using real-time PCR, antigen/antibody tests via ELISA and immunofluorescence assays as well as direct visualisation of blood parasites (e.g. malaria) using microscopy [1]. Staff from EMLab who were already trained to work with BSL3/4 level pathogens were deployed with the laboratories.

During 2014 and 2015 EMLab were deployed in Guinea, Nigeria, Liberia and Sierra Leone and provided more than 1,300 days of laboratory time [1]. Details on the 71 hours of training modules for preparation of mobile laboratory teams as well as complete list of components for the laboratory are available [1]. Sequencing laboratory modules were not part of the EMLab but a team from the UK with a sequencing laboratory in a suitcase joined the EMLab in Guinea [110, 111].

Figure 4. Schematic setup of the European Mobile Field Laboratory equipment and layout of two independent electrical supply lines in a tent or a fixed building. From [1].



5.3.2 *East African Community (EAC) Mobile Laboratory Network*

Overview. The EAC mobile laboratory (EAC-ML) network [112] is a regional network of mobile biosafety level (BSL) 3/4 laboratories that can be rapidly deployed to the location of disease outbreaks. During phase 1 of the project planning, nine modular laboratories were procured. Laboratory capabilities will be further enhanced in phase 2 of the project (upcoming) with six container mobile laboratories which will include BSL3 container laboratories equipped with bacterial culture facilities and whole-genome sequencing.

Location. Kenya, Burundi, Tanzania, Rwanda, Uganda and South Sudan (EAC countries). There will be one container per country [107].

Purpose. East Africa is prone to frequent disease outbreaks, which often happen in remote areas where BSL3/4 facilities are not available. Sample transport to national public health laboratories or overseas can significantly delay diagnosis and ongoing outbreak management. The mobile laboratories, which offer diagnostics of viral and bacterial diseases, bacterial culture and AMR profiling, have reduced sample turnaround time from days to an average of eight hours. Once in Phase 2 the laboratories will follow the WHO Global Antimicrobial Resistance Surveillance System (GLASS) guidelines and focus on AMR surveillance of the GLASS priority pathogens.

Modular laboratory set up. The EMLab was used as the blueprint for the EAC-ML. The modular laboratories can be set up in any existing infrastructure, e.g. a community centre, or in a tent if no structure is available. The laboratory equipment is kept in boxes that are transported, along with the laboratory personnel, in two Land Cruiser vehicles. The modular laboratories have a unilateral workflow with 12 workstations including donning and doffing area, sample reception, glovebox for sample inactivation, clean bench for reagent preparation, nucleic acid extraction bench, mastermix preparation bench, template addition bench, positive control bench, PCR workstation, ELISA workstation, office area and waste disposal/autoclaving area.

Key items of equipment. For the modular laboratories: Two −21°C portable compressor fridges/freezers and Va-Q-tainers; BioRad CFX96 PCR machines and Tecan Infinite ELISA

absorbance readers; the mobile gloveboxes (Koennecke, Berlin, Germany) in the laboratories provide BSL 3 and 4 facilities to contain and safely inactivate pathogens. Machines are mostly maintenance free and have few moving or sensitive parts, are resistant to the 0.5% bleach used in the laboratory and can operate in the hot and dusty conditions in the region.

The container laboratories will be 20-foot long extendable lorry containers that can be loaded onto trucks and transported to the field. They will contain: i) Laminar flow hood and incubators for bacterial culture, (ii) a Bactec blood culture system, (iii) MinION sequencing and bioinformatics facility for genotypic resistance testing, (iv) Nucleic acid extraction robot, PCR Hood, PCR machine for pathogen identification, (v) glovebox for inactivation of BSL3/4 pathogens, (vi) blood chemistry facility, (vi) Centrifuges, fridge/freezer, autoclaves, (vii) ICT and LIMS connectivity, (viii) back-up generator to assure autonomous operation of laboratory.

Training. The project had a training component combining practical laboratory sessions with an online e-learning platform and operational support training. Applying a 'Trainer of Trainers' concept, twelve regional trainers – two from each partner state – received in-depth 12-week training in 2018-19 in how to operate a mobile laboratory, biosafety and diagnostics of BSL3/4 pathogens. A mobile laboratory training facility was established at the EAC Secretariat based in Tanzania and several training activities took place there. In under a year the 12 regional trainers had taught a pool of 72 mobile laboratory operators. Seventeen different online training modules were also provided. The regional trainers and the International Air Transport Association, complemented by online resource, trained and certified 102 shipping staff in the region on safe sample transportation.

Key personnel. Modular laboratories: ideally six trained personnel are required to run the workstations [107]. Operational support: 18 drivers were taught maintenance of Toyota Land Cruisers. An additional 3-week training for biomedical engineers on mobile laboratory equipment was also provided to the drivers.

Energy/power demands. Modular laboratories: The power supply system can either integrate into existing grid power or run from a 3000W generator, with a backup uninterrupted power supply (UPS) system. The laboratories are equipped with a military-spec 12V lighting system that can be operated through the electrical system and UPS. The fridge/freezers can be powered through the 12V electrical system of the cars.

Costs. 10 million Euros in funding received from the German Development Bank to establish the nine modular laboratories in phase 1.

Data/information management. A LIMS is being installed in the mobile laboratories and training offered.

Key outcomes. In phase 1 of the network the project demonstrated a reduction from days to 8 hours for sample turnaround times. This had an immediate impact on the time individual patients spent in isolation wards (for viral haemorrhagic fevers) and also contributed to interrupting chains of transmission more quickly. The mobile laboratories have taken part in outbreak responses against Dengue, Ebola and COVID-19.

Other considerations. Five key areas were identified by the project leads to support the development of a network such as the EAC-ML [107]:

1. Transferring ownership to local partners through regional project coordination
2. Procurement of equipment and consumables involving regional suppliers
3. Technical aspects of mobile laboratory design
4. Training of local laboratory operators
5. Field simulation exercises and participation in outbreak responses.

5.4 Mobile laboratories and sequencing

Genome sequencing carried out by mobile laboratories provides significant opportunities in terms of real time epidemiological assessment and response during outbreaks. Information gained from sequencing allows: pathogen characterisation; understanding of pathogen and its evolution including potential new variants; identification of genetic changes that could alter diagnostic qPCR/PCR targets or the effectiveness of vaccines and therapeutics; and genomic surveillance. While genome sequencing requires specialised analyses and can be helpful in providing additional data to manage an outbreak, efforts are needed to ensure that the normal functioning of the mobile diagnostic laboratory and its ability to provide quick turnaround times on patient samples is not disrupted by the addition of sequencing capability [102].

One of the main challenges in providing sequencing in a mobile laboratory setting is the equipment and how it performs in more challenging settings. As outlined in Chapter 3, sequencing equipment can be highly sensitive with carefully aligned instrumentation which often requires specialist engineers to install, move and service, and to ensure it is working correctly. Many manufacturers can nullify guarantees and warranties on their equipment if it is moved and handled by someone other than their own service engineers. In addition, many sequencing platforms are large in size (Tables 5 and 6) and need a consistent source of power to protect them from unexpected electricity surges. The equipment must also be protected from the elements, i.e. dust, rain and extreme heat and humidity, to ensure their proper functioning and to prevent breakdowns. For these reasons, and as highlighted in Chapter 3, few of the currently available sequencing machines are suitable for mobile laboratories. Many of the key considerations for mobile laboratories are particularly pertinent for mobile sequencing laboratories and include [113]:

- Portability. No more than two people should be able to transport or carry the ‘laboratory’ (although this is not the case for some of the larger container laboratories)
- Robustness. The laboratory equipment and protocols should be able to handle harsh environments such as high humidity, temperatures, dust and pests
- Rapidity. Laboratory set up to enable sequencing needs to take place within hours, not days. The sequencing workflows also need to be rapid (hours, not days)
- Field suitability. The laboratory needs to be able to operate in the absence of air-conditioning, cold chain, or internet
- Simplicity and ease of training. The set up and running of the sequencing workflows should be easy to teach and operate, so that local staff can ensure sustained use for future outbreaks
- Accuracy and interpretability. The sequencing data needs to be accurate, reliable and with robust reporting strategies
- Funding. Sustainable sources are needed to ensure ongoing operations.

Sequencing laboratory workflow. The laboratory examples listed above and in Table 7 (below) follow the same basic principles in terms of sequencing workflows (Section 4.2). These workflows require up to three distinct workstations that are clearly designated for specific tasks and aim to minimise contamination:

- Sample preparation and extraction
- Mastermix preparation
- Sequencing preparation – where the mastermix and samples are combined for sequencing.

If dealing with a potentially infectious agent there may also be the requirement of a ‘glove box’ workstation to inactivate the sample before preparation and extraction. The ‘mastermix’ station is where the consumables are mixed and prepared and is considered a ‘clean’ station as no samples, DNA, RNA, or amplified product should ever be at or near these stations.

This is to avoid contamination of the mastermix and reagents used to sequence the samples. The final station is where the mastermix and extracted samples are combined and prepared for sequencing, although some may do this where the sample extraction has taken place. The final step is to load the prepared samples onto the sequencing equipment, which should be located at the last station or nearby.

The need to avoid contamination when doing molecular biology techniques means that in static laboratories these workstations will often be in separate rooms. For example, clean rooms can often be at the other end of a corridor to the DNA extraction room. Some laboratories will also have protocols in place outlining that different laboratory coats are used at each workstation, or that staff that have worked in the DNA extraction room are not allowed to enter the mastermix/clean room in the same day. The importance of this workflow should be carefully considered when setting up a mobile laboratory, particularly one where equipment is 'unpacked' and set up. Demonstrations of these different workstations can be seen in the following publications: [1, 107, 114].

5.4.1 Laboratory for SARS-CoV-2 testing and sequencing in a shipping container

Overview. Founded in 2018, OpenCell.bio provides laboratories for biotechnology startups. In response to the COVID-19 pandemic a multidisciplinary team (including UK university researchers) led by OpenCell.bio developed a low-cost, rapidly deployable COVID-19 testing laboratory, and a sequencing laboratory, inside shipping containers.

Location. UK: London and the Bailiwick of Jersey, a geographically remote island in the English Channel.

Purpose. The COVID-19 pandemic has challenged diagnostic systems globally and attempts to scale-up diagnostic testing have been hampered by limited laboratory infrastructure and logistical challenges. OpenCell developed a BSL-2 container laboratory optimised for automated COVID-19 diagnostics (qPCR) [115] and another container laboratory containing portable sequencing technology [116] to support population sampling, the delivery of diagnostic testing and sequencing in areas with limited laboratory infrastructure. Schematics of the container layout are available in the preprint publications [115, 116]. The team demonstrate a reproducible workflow for RT-qPCR COVID-19 testing as well as sequencing, using open-source hardware and reagents.

Laboratory structure. The laboratories are standard 40-foot long, 'high' shipping containers, adapted to contain a BSL 2+ laboratory that meets the ISO15189 standard for a medical laboratory. Adaptations to achieve this include interior surfaces made of hygienic and easily sterilisable materials and all modifications ensure that the containers maintain compliance with international sea-freight standards and can still be shipped as normal freight with the original Container Safety Convention plates. This permits international transport of the laboratory to any major port. Laboratory units can be transported by truck to any location and can be placed on a tarmacked space, such as a car park or with minor additional groundwork on uneven ground. The production time of a container laboratory, including the interior, takes two working days. The container laboratories are designed to plug into mains electricity supply and water systems of a nearby building or to run from a water tank and generator. The water system is designed to prevent contamination of freshwater through an anti-backflow valve and to store wastewater for easy removal by clinical waste disposal services alongside other physical waste.

Laboratory workflow. RT-qPCR testing for COVID-19 requires a linear physical layout matching the workflow and cleanliness requirements of the assay. The workspace is divided into three different sections [115]:

- Station A: Unpackaging and computational logging of samples
- Station B: RNA extraction
- Station C: RT-qPCR

Station C is physically separated from Station A and B but also subdivided further, with a heavy-duty self-sealing PVC curtain insulating the bench used to prepare plates for PCR from the qPCR machines. With this linear physical layout, personnel and samples move from Station A to Station C, changing PPE from B to C, with minimal risk of personal and sample or nucleic acid contamination.

The separate sequencing container laboratory also requires a linear workflow, divided into three stations [116]:

- Station A: RNA extraction from samples, reverse transcription of RNA to cDNA. Library preparation, PCR plate preparation.
- Station B: Sequencing
- Station C: PCR for cDNA synthesis

Key items of equipment. RT-qPCR: OpenCell have developed automated protocols for RNA extraction and qPCR plate preparation.

- Station A: Contains a BSL2+ microbial safety cabinet for sample inactivation of active SARS-CoV-2, sink, autoclave, PC with OpenCell LIMS, storage for items not requiring refrigeration, clinical waste bin
- Station B: Contains four Opentrons (OT2) liquid handlers, to conduct multiple simultaneous RNA extractions on samples. The RNA extraction process employs the open-source Bio-On-Magnetic-Bead (BOMB) protocol, which uses DIY reagents and is amenable to automation. The OT2 liquid handling robot has a low initial cost relative to other liquid handling robots, can utilise generic plasticware and run protocols written in the coding language Python. Fridge, storage unit, laptop.
- Station C: Area 1 contains one OT2 for preparing plates for RT-qPCR, storage unit, laptop; Area 2 contains up to two qPCR machines, freezer, storage unit. A clinical waste bin is provided at the exit.

Sequencing: 154 samples were prepared and sequenced using ARTIC network protocols.

- Station A: Two safety cabinets to avoid contamination – one for cDNA mastermix, the other for viral RNA. RNA extraction used the MagMAX™ Express-96 Deep Well Magnetic Particle Processor, SARS-CoV-2 virus detection performed using 2019-nCoV RealStar SARS-CoV-2 RT-qPCR Kit (Altona Diagnostics) by RT-qPCR assay (Quantstudio™ 2, Applied Biosystems). Quantus Fluorometer to quantify barcoded amplicons. Laptop.
- Station B: Oxford Nanopore Technologies MinIon sequencer. High performance laptop.
- Station C: PCR machines. Laptop.

The use of open-source hardware and reagents allow laboratories to navigate around limited supply chains.

Training. While most RT-qPCR testing protocols rely on expert operators, the semi-automated OpenCell protocol, using the OT2 robot, may theoretically be run by a technician with generic laboratory experience. No information provided on sequencing laboratory training.

Key personnel. The diagnostic laboratory requires six members of staff working in shifts [117]. Each qPCR diagnostic laboratory is capable of running up to 2,400 tests in 24 hours. A protocol was developed and tested in a minimal set-up (one qPCR, two OT2s) which allowed for 1000 tests per day. No information provided on sequencing laboratory personnel.

Energy/power demands. Containers are designed to be easily plugged into the mains and water system of a nearby building or can be run via water tank and generator. Electricity:

Commando socket on the outside for the electricity 16amp/connector to connect to mains, plus a back-up generator.

Costs. Not provided. The Government of Jersey contract with OpenCell ended in November 2021 when a new hospital laboratory able to perform PCR tests opened [118].

Data/information management. An open-source LIMS was used and included QR scanning done via webcams or phones. To ensure data accessibility for the public health authorities or wider medical bodies the team investigated the HL7 Fast Healthcare Interoperability Resources (FHIR) protocols which are in widespread NHS use [119, 120]. An extension of this software could be to run to a FHIR compatible proxy server between the LIMS and outside providers allowing results to be securely communicated, enabling data integration with the NHS or any other international healthcare system architectures.

Key outcomes. A two-step validation process for the qPCR automated assay laboratory was done, first by running samples spiked with synthetic SARS-CoV-2 genomic RNA to optimise their process and subsequently on patient samples to validate their assay and workflow. Whilst operational on Jersey 598,155 RT-qPCR tests were run, identifying 8,950 positive cases with sequencing of 154 samples [116].

Open-source hardware, code, and reagents permits the use of generic plastic ware and easily obtained reagents, circumventing supply chain issues by adapting to local or transient shortages. The shipping container laboratory permits global transport and distribution. A custom open-source data management system tracks samples through the system, using the FHIR protocol, an international standard which enables integration into the UK National Health Service or healthcare systems in other countries.

Protocol changes were made that reduced risk of cross-contamination as well as simplified the process by removing the need for a centrifuge and a fume hood for volatile toxic compounds. Other changes reduced the number of pipetting steps, saving time and reducing potential errors and consumables needed. Continuing efforts are underway to identify alternative protocols and improve the efficiencies which would save on plastic ware, time and reagents.

Other considerations. The flexibility of OpenCell's system allows hospitals to integrate the unit with their own testing process, existing reagents, waste management system, and/or workforce. The addition of a dedicated, separate facility enables a hospital to increase the scale of its testing without impacting other essential in-house laboratory work and maintains the safety of the main hospital building.

5.5 Sequencing only mobile laboratories

In the past decade, with the development of smaller devices and improved engineering, it has become possible to carry out sequencing outside of traditional laboratory infrastructure. Sequencing has been deployed as part of a larger mobile laboratory responses or has been the sole purpose of the laboratory. Mobile laboratories demonstrate that it is feasible for sequencing to function in these conditions. A variety of options for the set-up of mobile sequencing laboratories have been used, these include:

- The 'sequencing lab in a suitcase' where no set up is required [60, 121, 122]
- Sequencing laboratories packed into suitcases or bags that can be carried by a single person and unpacked where it is needed, such as on a desk in an office, school, in a clinic laboratory or the back of a car [2, 108, 109, 123]
- Sequencing laboratory installed in a vehicle – could be a car, trailer, caravan or small truck [124-127]
- Sequencing laboratory in a container which can be transported on the back of a large truck [112, 116]
- Sequencing laboratory transported in boxes to a field site and installed in tents or structures such as containers or buildings [2, 59, 109].

With sequencing technology improving and the availability of smaller devices, sequencing has been used in a mobile laboratory setting, almost exclusively using Oxford Nanopore Technologies sequencing platforms (Table 7).

Table 7. Examples of sequencing carried out in mobile laboratory settings.

Laboratory for sequencing in a:	Pathogens under investigation	Sequencing equipment	Regions / countries used	Reference
Suitcase (no need to unpack)	Avian influenza, Ebola	MinION	Guinea, West Africa	[60, 121, 122]
Suitcase and unpacked in the back of a car or on a desk	Rabies	MinION	Kenya, Tanzania, the Philippines, Guatemala, India, and Vietnam	[109]
Suitcase and unpacked almost anywhere, most in buildings	Ebola, necrotic skin infection, SARS-CoV-2, rabies	MinION	Guinea, Republic of Sao Tome, UK	[2, 108]
Land Rover	Malaria	MinION	Namibia, Zambia, Tanzania and Kenya	[124, 125]
Caravan	Zika	MinION	Latin America	[126]
Trailer / truck	SARS-CoV-2	MinION and GridION (LampORE)	UK	[127]
Shipping container	SARS-CoV-2	MinION	Bailiwick of Jersey, UK (one of the UK Channel Islands)	[116]
Field tent	Ebola	Ion Torrent Chef and PGM	Sierra Leone	[59]

5.5.1 MicroLabs based genomic surveillance of SARS-CoV-2 , India

Overview. The aim was to bring SARS-CoV-2 sequencing capacity close to sites where samples were being collected and to focus on high-priority sequencing in resource limited settings, with minimal infrastructural requirements and turnaround times. This will help monitor the introduction of variants of concern, evolution of sub-lineages and support the ongoing public health planning.

Locations.

- Maharshi Dayanand University (MDU), Rohtak, Haryana. Chosen to focus on an area with a rural population as catchment, in close coordination with the district civil hospital
- CSIR-Central Drug Research Institute (CSIR-CDRI), Lucknow, Uttar Pradesh. Aim to cater to the biggest state of India, in terms of population size
- CSIR-North East Institute of Science and Technology (CSIR-NEIST), Jorhat, Assam. Chosen to be the regional centre for SARS-CoV-2 sequencing for the whole North Eastern region of India.

Purpose. To provide a successful template for future pandemic preparedness in India with genomic surveillance as a core function.

To aid and augment SARS-CoV-2 genomic surveillance through:

- Discovery of variants of interest (VOI) and variants of concern (VOC)

- Tracking the evolution of VOIs to VOCs
- Identifying genomic hotspots based on WGS
- Identifying mutations associated with disease severity for pre-emptive surveillance in targeted geographical locations.

Laboratory structure: A simple design capable of being set up in either a mobile container van or in a room with aluminium temporary partitions between the three areas.

Laboratory workflow. Three segregated areas for:

- Section 1: sample preparation/isolation of viral nucleic acid
- Section 2: pre-PCR/library preparation,
- Section 3: post-PCR step/sequencing and data analysis.

An experimental and computational pipeline, optimised by the CSIR-Institute of Genomics and Integrative Biology (IGIB) is being used in alignment with global standards.

Key items of equipment. general equipment as needed per section, including: pipettes, spinner, vortex.

- Section 1: Air conditioning (AC), BSL hood, DNA/RNA quantification (Qubit), UPS, autoclave
- Section 2: AC, 4°C fridge and -20°C freezer, PCR machine, centrifuge, magnetic stand
- Section 3: AC, computer, data storage, Oxford Nanopore Technologies MinION sequencer.

Training. Training is integral to the success of the MicroLab based genomic surveillance for near patient sequencing and requires regular updates from the experimental and the computational perspective. IGIB carries out monthly meetings with the MicroLab PIs to ensure key learning is shared in quickly and any queries are also answered swiftly. In addition, training for the new manpower joining the group is also equally important.

Key personnel. Laboratory technicians to carry out experimental work and bioinformaticians to provide computational manpower, public health officials, clinical partnerships, funding support partnership, and global data sharing facilitators.

Energy/power demands. None of the instruments mentioned above have high energy demands.

Data/information management. Measures have been put in place to:

- Ensure data homogeneity to support seamless sharing and learning
- Reduce chances of mistakes while data sharing and drawing inference/s
- Allow safe data storage and archiving.

Key Outcomes. First of its kind MicroLab set up in a rural location for genomics based surveillance. This will be important for the future pandemic preparedness, especially future COVID wave/s and eventually other diseases of continuing concern – tuberculosis, Dengue, and human papilloma virus. A locally developed bioinformatics tool has supported data analyses. For proactive surveillance of the biodiverse North Eastern region of India, in addition to the present SARS-CoV-2 genomic surveillance, the developed infrastructure and expertise can be used for future pathogen discovery.

Other considerations. Engagement and continued learning has been key to adapting to emerging and evolving needs.

5.5.2 Considerations for mobile sequencing laboratories

Equipment and consumable requirements for sequencing. Costs of establishing and running mobile sequencing laboratories are context dependent and vary through time and geographical location, and it is sometimes unclear how many samples can be analysed given the consumables used. However, a number of sequencing projects have provided equipment lists with some estimates of costs for their mobile laboratories:

- The laboratory in a suitcase used during the Ebola outbreak is detailed on the ARTIC website and the related publication detail equipment used and provide estimated costs [2, 128].
- Portable sequencing done in Peru as a teaching tool in conservation and biodiversity research has an equipment list with estimated costs [129].
- The laboratory in a suitcase used to investigate rabies in six LMICs is based on the ARTIC sequencing protocols with details of equipment on github and protocols.io [109].
- Zika Virus Amplification Using Strand Displacement Isothermal Method and Sequencing Using Nanopore Technology is a chapter in the book *Zika Virus Methods and Protocols,* part of the *Methods in Molecular Biology* book series [88].

Bias in reporting. A common observation in the literature and confirmed by experts has been that only successfully implemented mobile sequencing laboratories have been reported and that there is limited or no information on failed attempts. While the reasons for failure are not being captured, we can only learn from situations where implementation has been successful. There was also high variability in the information provided between the different examples found. There is no consistency in the reporting of the details of mobile laboratories used, which present difficulties when comparing different mobile laboratory approaches.

5.6 Conclusions

The feasibility of mobile laboratories, including in field settings, has been demonstrated by the examples outlined. Sequencing technologies have also been successfully deployed in mobile laboratories, demonstrating that the technology can work in more challenging conditions and provide valuable support to authorities managing outbreaks in a number of countries. There continues to be a need to improve the sequencing technology, infrastructure and protocol standardisation, and to provide training in the skills required to operate such laboratories. As the technology improves there are likely to be further opportunities to expand the use of sequencing in mobile and field laboratories.

6 Near patient sequencing case studies

6.1 Sepsis

Sepsis is defined as 'life-threatening organ dysfunction, caused by a dysregulated host immune response to infection' [130]. Infections that result in sepsis can be acquired in both community and healthcare settings. Bacterial infections are the most common cause of sepsis, but the prevalence of different sepsis causing pathogens will also vary by region. For example, in the Southern Hemisphere and tropical regions sepsis can also arise from viral diseases such as measles and viral haemorrhagic fevers, and protozoal diseases such as malaria [131].

In 2017, an estimated 48.9 million cases and 11 million sepsis-related deaths were recorded worldwide, which corresponds to 19.7% of all global deaths [132]. As most systematic data collection in sepsis is carried out in HICs and very little in LMICs, it is difficult to estimate the true incidence and mortality of sepsis in LMICs [133]. However, 80% of the global mortality attributed to infections is from LMICs, so it is assumed that they also suffer a disproportionately high burden of sepsis [131]. It has been estimated that in HICs, whilst sepsis is fatal in around 30-40% of cases, this rises to up to 80% in the most resource limited settings [131]. Those at higher risk of developing and dying from sepsis include the elderly (particularly those with comorbidities affecting immunity) [134], pregnant and post-partum women, neonates and young children. In 2017, it was estimated almost half of all global sepsis cases occurred in children, with an estimated 20 million cases and 2.9 million deaths occurring in children under five years old [132].

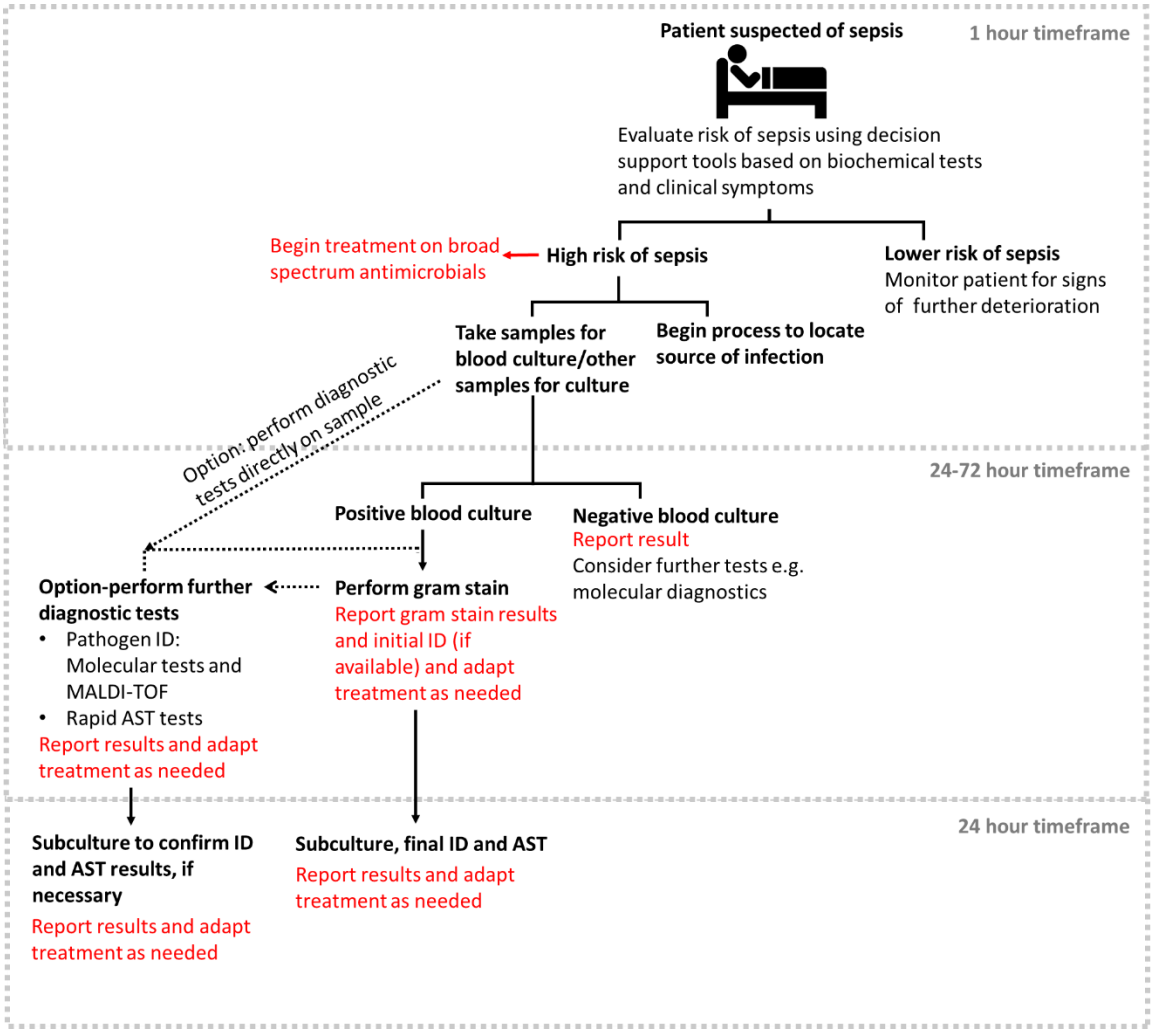
6.1.1 Current pathways for sepsis diagnosis and treatment

Figure 5 outlines key steps that take place in sepsis management, along with approximate timespans for each step. The priority in sepsis management is the initial recognition of possible sepsis symptoms and deciding if a patient is at high, moderate or low risk of sepsis. A variety of decision support tools and physiological tests are used for this purpose. The aim is to provide treatment, including broad spectrum antimicrobials, to higher risk patients whilst lower risk patients may be monitored for further signs of sepsis. It is recommended that treatment commences within one hour of the recognition of the initial signs of sepsis, in order to increase chances of survival and a full recovery [135]. It is less urgent but also important to identify the source in the body from which the infection originates. For example, for sepsis occurring in ICU settings, the expert consensus recommendations group of the European Society of Intensive Care Medicine and the Mahidol-Oxford Research Unit in Bangkok, Thailand suggest that source identification and control should be performed within 12 hours of admission to hospital [135].

The next key step is to identify the pathogen causing the infection and if it possesses antimicrobial resistance, so that the broad-spectrum antimicrobial therapy can be deescalated (the spectrum narrowed) if possible or the therapy can be altered to make it more suitable for the pathogen detected [136]. This is important to ensure that the patient receives the most effective therapy for their condition, for example to prevent the prescription of antibiotics for a viral infection. In addition, whilst broad-spectrum antimicrobials are important in treating sepsis, antimicrobials themselves are associated with adverse events as well as complications due to AMR, making it important to avoid unnecessary use [137, 138]. There is no defined time period over which antibiotic therapy should be deescalated or modified based on culture results, and more research is needed on the benefits. However, organisations such as the Surviving Sepsis Campaign encourage that it is performed as soon as possible, due to the individual and societal risks associated with unnecessary use of antibiotics [139].

Most sepsis is caused by bacterial infection, and so it is typically recommended that blood culture be performed in order to identify the pathogen [140]. Culture of other samples relevant to the infection may also be used, for example urine and cerebrospinal fluid [140, 141]. Ideally samples for culture should be taken before the patient is treated with antimicrobials. Positive cultures then undergo subculture and antibiotic susceptibility testing (AST). Alternative and/or complementary techniques to culture exist, including further molecular tests such as the BIOFIRE Blood Culture Identification 2 Panel (BCID2) [142], or MALDI-TOF mass spectrometry based tests [143]. These tests can help more accurately and rapidly identify pathogens, often using positive culture media or mixed colonies so avoiding the need for further subculture. Whilst these methods can also provide antimicrobial resistance information useful for treatment decision making, currently further culture is required if a full antimicrobial susceptibility profile is required. A few molecular methods are available that bypass the use of culture altogether for pathogen identification, some also generating AMR information. These include PCR based methods such as the T2Bacteria Panel [144], as well as NGS based methods. These methods can be challenging to develop and apply due to low levels of pathogens in clinical samples but can deliver faster results and be necessary if non-bacterial infection is suspected, as pathogens such as viruses and fungi are often less amenable to culture.

Figure 5. Key steps in sepsis management in a HIC setting, with approximate timeframes for each stage if a culture based method of pathogen identification is used. Information based on UK guidelines: [140].



6.1.2 Challenges to sepsis management in lower resource settings

There are several challenges with current methods of sepsis management, some of them relevant to all settings and others more specific to lower resource settings.

Difficulties with the initial recognition of sepsis

Sepsis is a syndrome consisting of an assortment of what can be vague symptoms for which no one diagnostic test currently exists, which can make it hard to identify. As most guidelines for sepsis recognition have been developed in HICs, they may not be as relevant for LMICs. This can be due to differences in the causes of sepsis resulting in different symptoms and clinical indications of sepsis. In addition, LMICs may have limited access to some of the tests used to help diagnose sepsis, and test performance may not have been validated in the relevant populations. Factors affecting test performance can include higher background levels of inflammation in some LMIC populations, which could interfere with some host biomarker-based tests. In low and high resource settings, tests to rapidly identify infections, as well as distinguish between viral and bacterial infections, are required. This would guide more appropriate use of empirical antimicrobials, to increase treatment efficacy and reduce AMR [145]. MeMedBV is one example of a test that has been FDA approved, however its applicability to LMICs remains unknown [146]. Any new tests being considered should also be suitable for point of care use within the one-hour timeframe needed to diagnose sepsis and initiate treatment.

Relevance of empirical antimicrobials and guidelines for initial sepsis treatment

The empirical broad-spectrum antimicrobials used after initial detection of sepsis should be appropriate for the types of microbes prevalent in a particular region, including taking into account the levels of AMR within the microbial population. However, regional data of sepsis causing pathogens is needed to inform antimicrobial use and other treatment guidelines, and this is often lacking in LMICs [131]. As a result, treatment guidelines are based on data from higher income countries, which may not be suitable for lower income settings. For example, as part of the BARNARDS study, the efficacy of antibiotic combinations commonly used in LMICs for neonatal sepsis was assessed. The study found that gram-negative isolates were ‘overwhelmingly resistant’ to ampicillin (379 [97·2%] of 390 tested) and to gentamicin (274 [70·3%] of 390 tested) [147]. Despite this, WHO guidance recommends ampicillin-gentamicin for empirical use in neonatal sepsis, and in the BARNARDS cohort ampicillin–gentamicin was administered by almost all sites across Africa and in Bangladesh. Whilst ampicillin-gentamicin is a relatively cheap antibiotic combination and may be effective in HICs, the study suggests its use in LMICs should be reviewed. There have been calls for more local sepsis-surveillance efforts in LMICs, in order to create more relevant regional sepsis management guidelines [135, 148]. This is not a problem unique to sepsis, but in the diagnosis of many infectious diseases with multiple overlapping symptoms, e.g. febrile disease [55].

Limitations in use of culture for pathogen identity and AST

As described in Figure 5, the most common method for identifying sepsis-causing pathogens is via culture, which is typically considered the ‘gold-standard’ or reference method for diagnosis of bloodstream infections [139, 149]. Subculture is then used to perform antimicrobial susceptibility testing. However, blood culture (the most commonly used sample) is estimated to detect the causative organism in only 15-30% of septic patients, with culture of other samples producing a positive result in 20-30% of patients. Therefore for at least 40% of patients culture yields no results [150]. This could be for several reasons, including lack of sensitivity of blood culture systems to detect pathogen growth, for reasons such as insufficient sample volume and/or use of antibiotics prior to culture limiting the amount of pathogen(s) present in a sample. It is also increasingly recognised that a septic infection can remain localised and does not have to result in bloodstream infection, therefore blood culture will not be able to detect all cases of sepsis [149]. Furthermore, sepsis can be

caused by pathogens not amenable to the culture systems used. In addition, culture has several other intrinsic limitations:

- For pathogens not amenable to culture – e.g. some viruses and fungi –pathogen detection is not possible and further culture-independent tests are required
- There is a risk of contamination during sample collection for culture, estimated to be as high as 10% in some HICs and likely higher in LMICs [151]
- For neonates and paediatric populations, it is challenging to provide sufficient sample volume for successful culture
- Culture-based methods can be slow, on average taking from 24-72 hours from sample receipt to pathogen identification, during which time the patient may continue to receive ineffective or unnecessary broad spectrum antimicrobials, rather than a therapy tailored to their pathogen. In LMICs, automated systems are rarely available, and the manual culture used instead can further increase time to pathogen identification [151]
- Further subculture and AST are needed to provide information on AMR, taking at least another 24 hours.

Despite these limitations, culture is widely used in HICs, which typically have appropriate culture facilities available (e.g. in UK typically containment level 2 unless a hazard group 3 organism is suspected, requiring containment level 3 [140]), with the required equipment and materials including automated culture systems, as well as the staff and expertise needed to perform culture and assess results. However, in LMICs financial, logistical and infrastructural constraints mean culture is often not performed [54, 151].

6.1.3 Potential of near-patient NGS as a sepsis diagnostic test

NGS tests being developed for sepsis diagnostics fall into two main categories; metagenomic NGS (mNGS) tests which sequence all genomic material in a sample, and pathogen agnostic targeted tests which sequence conserved targets shared across types of microorganisms such as the 16S rRNA gene in bacteria and the 18S rRNA gene in fungi, allowing multiple pathogens to be identified. Pathogen agnostic targeted tests may also incorporate specific targets such as AMR or viral genes. These methods, along with their key characteristics, are discussed in more detail in Chapter 4.

NGS based tests would ideally be used to identify sepsis-causing pathogens directly from clinical samples, to overcome some of the limitations of culture described above. In addition, NGS is being explored as a method for pathogen identification following initial positive culture results but using culture broth as the sample rather than culture isolates. In this case it could act as a more rapid but still culture dependent method, as an alternative to other molecular pathogen identification techniques such as MALDI-TOF and PCR panels [152]. As the shortest possible turnaround time is required for sepsis diagnostic tests, ideally NGS based tests would be performed in near-patient settings such as hospital laboratories, or even outside of a laboratory environment.

The first WHO report on the global epidemiology and burden of sepsis lists several requirements for an ideal diagnostic tests for sepsis-causing infections [153]. Most NGS tests for sepsis are still in an early stage of development and so their utility and feasibility as a sepsis diagnostic tool is currently uncertain. The future potential of these tests to fulfil the WHO requirements is outlined below:

Requirement 1: Rapidly identify pathogens broadly (bacteria, virus, parasite, fungi). mNGS tests which sequence all material in a sample are capable of identifying and classifying any type of pathogen present at detectable levels in a sample, as long as a reference genome sequence exists. NGS tests are unlikely to provide the rapid results needed to make decisions on broad spectrum antimicrobials, where results are required within one hour of sepsis diagnosis. However, NGS tests have the potential to substantially speed up the process of pathogen identification following a diagnosis of sepsis. For

example, in a small proof-of concept study, a mNGS assay using the ONT minion was estimated to take six hours from sample receipt to identification of the sepsis-causing pathogen [26]. This is faster than culture, which can take several days to produce a pathogen identification. Another small study using a targeted amplification of the 16s rRNA gene for bacteria, the IST1/2 gene for fungi, and the rpoB gene for Mycobacterium spp. had an average turnaround time of around 6-18 hours [154]. Other methods based on Illumina sequencing are slower but can produce pathogen identification within 24 hours in research settings [155]. The length of time taken in real-world clinical settings requires further evaluation for all sequencing methods.

Requirement 2: Be highly sensitive and specific so as to guide antimicrobial therapy, limit antibiotic overuse, and inhibit AMR development. The sensitivity and specificity of NGS tests varies depending on factors such as the type of sample sequenced (e.g. blood, urine or a culture broth) and the NGS method used (e.g. untargeted or targeted sequencing). Direct from blood sample NGS appears to be more sensitive for pathogen identification than use of blood culture. For example, a study in Germany using Illumina mNGS to analyse the blood plasma of 48 septic patients found that at sepsis onset, blood culture was able to detect 33% of pathogens whilst mNGS was able to identify 72%. The authors stated that 96% percent of the positive sequencing results were plausible and would have led to a change to a more adequate therapy in 53% of cases as assessed by an independent expert panel [155]. In a different small retrospective POC study using targeted sequencing of the 16s rRNA gene for bacteria, the IST1/2 gene for fungi, and the rpoB gene for Mycobacterium spp. was able to identify a pathogenic or likely pathogenic microorganisms in all 11 cases, compared to culture where a pathogen was identified in only two cases [154].

An expert interviewee cautioned that although blood cultures are commonly used for sepsis, part of the reason that only on average 40% of blood cultures yield results is because sepsis-causing infections can remain localised, and so pathogens will not enter or be detected in the bloodstream. Therefore, whilst NGS of blood samples may be more sensitive than blood culture, it will not be possible to diagnose pathogens in cases where they are not found in the bloodstream. However, methods that sequence cell-free pathogen DNA in blood plasma may be able to help address this issue, as cell-free DNA is released from locations throughout the body. Due to the ability of pathogen agnostic methods in particular to identify all microbes in a sample, NGS can also be less specific than culture (i.e. there are a higher number of false positive results), as it may more readily detect commensal and/or contaminating microorganisms that are not the cause of infection.

Requirement 3: Use readily available clinical samples (for example, whole blood) that do not require processing or culture. Most NGS tests in use or development for sepsis diagnostics are designed for use directly on clinical samples without the need for culture (see Tables 8 and 9). This is in contrast to molecular PCR panels available for bloodstream pathogen detection such as the BioFire BCID2 Panel, which requires culture broth [142]. The level of processing required to prepare samples for sequencing may still be a limiting factor in terms of the sample and library preparation required as part of the sequencing workflow (see Chapter 4). However, there may be the ability to automate all or part of the process in the future.

Requirement 4: Allow detection of multiple pathogens simultaneously. NGS methods to sequence samples directly allow the detection of multiple pathogens simultaneously. Pathogen agnostic mNGS methods have the potential to identify all known pathogens in a sample, by sequencing all genomic material present in a sample and comparing the sequence results to databases to determine the identity of known pathogens. As such, assays based on these methods offer the most potential to detect all types of pathogens associated with sepsis. This could be particularly useful in settings where there is a lack of information on the most likely sepsis-causing pathogens making pathogen targeted tests less suitable. It could be useful in scenarios where both viral and bacterial pathogens may be implicated in sepsis, as mNGS can be used to detect both types of pathogens. Examples

of mNGS assays being evaluated for sepsis pathogen identification in clinical trials are shown in Table 8. Similarly, sequencing of targets conserved across classes of pathogens such as the 16S rRNA gene in bacteria and the 18S rRNA gene in fungi can be used to allow sequencing to be performed in a pathogen agnostic manner, and are technologies being developed by some commercial providers (see Table 9). However this approach is not suitable for detection of all viral pathogens, although they can be combined with virus specific targets if required. In contrast the BioFire BCID2 PCR Panel can only be used for a limited number of pre-defined targets, 33 targets to support pathogen identification and 10 AMR gene targets [142].

Requirement 5: Detect drug susceptibility and resistance. As well as determining pathogen identification, the ability of NGS to predict drug susceptibility and resistance based on genetic sequences present is also being explored, by using NGS to detect genetic features such as AMR associated genes. This can allow pathogen identification and AMR prediction to be performed within the same test, facilitating more rapid detection of resistance. Currently phenotypic methods are still required to confirm predictions of drug susceptibility, however NGS may be used alone in future.

Requirement 6: Be simple to use with minimum training required. As for other clinical applications, most NGS testing being developed for sepsis currently requires trained laboratory staff to perform tests and analyse sequence data to deliver results. mNGS in particular is a complicated to perform and analyse. However, there is potential for tests to become more automated and user friendly in future. For example, this is the intention for the AutoSepT and LiDia-Seq tests described in Table 9.

Requirement 7: Be relatively low cost. NGS tests are typically considered a relatively expensive form of technology, however the cost of NGS is gradually decreasing. Use of NGS tests may also improve patient outcomes and use of antimicrobials, provide more detailed data for surveillance purposes as well as diagnostics, and require less infrastructure than other methods. Therefore the cost-effectiveness of NGS for different clinical needs will require evaluation, and specifically for lower resource contexts. Cost may also vary considerably depending on sequencing equipment used and current laboratory infrastructure. For example, cost per sample in a study using mNGS on an ONT nanopore, including the sample and library preparation, was US$1,100. This compares to US$300 per sample if Illumina sequencing was used for the same purpose [26].

6.1.4 Current development status of near-patient NGS as a sepsis diagnostic test

NGS based tests designed for sepsis pathogen identification for routine diagnostic purposes are still mostly in the research phase and are not yet routinely used even in high income countries. mNGS methods in particular are challenging to develop due to high levels of host DNA reducing the signal to noise ratio; other limitations are described in Chapter 4. Despite this, several prospective clinical trials described in Table 8 are currently underway exploring the use of NGS in the diagnosis of bacterial sepsis in critically ill adults and children, all in higher income countries. These clinical trials are not evaluating NGS performed in near patient settings, all sequencing is taking place centrally. However, evidence from these trials will help to establish if sequencing provides clinical utility in improving patient outcomes over standard culture-based diagnostics. Whether the technologies used are then applicable to near patient settings will depend on the available resources of the healthcare facilities. Further work may be needed to develop workflows and validate the sequencing assays in more resource-limited settings.

Several commercial services that perform NGS based tests on samples from sepsis patients shipped to a central laboratory do exist, highlighting the feasibility of using this method for pathogen identification. However these are only used in specific situations, and are not commonly used even in HICs. Other companies are developing tests designed specifically

for near patient settings, with end-to-end workflows. Examples of some key commercial tests available and in development are shown in Table 9.

Table 8. Current prospective trials evaluating use of NGS as pathogen diagnostic tool in sepsis

Trial name	Country/year registered	Trial type	Summary	Sequencing approach	Sequencing technology
The Next-Generation Sequencing Diagnostics of Bacteremia in Sepsis (NextGeneSiS) study [156]	Germany, 2017	observational, prospective trial, non-interventional, multicentre	Aims to assess the diagnostic performance of sequencing in a cohort of 500 patients with sepsis, compared to culture-based methods [9]. Will use a panel of independent clinical specialists to retrospectively identify potential changes in patient management that would have been made based on the results.	mNGS analysis of cell free DNA in blood plasma samples	Samples sequenced centrally using mNGS on an Illumina HiSeq2500; not performed in a near-patient context.
The Next GeneSiPS-Trial [157]	Germany, 2018	observational, prospective trial, non-interventional, multicentre	This trial extends the NextGeneSiS trial to assess the diagnostic performance of an NGS-based approach for the identification of causative pathogens in critically ill children of different ages (150 in total) with suspected or proven severe sepsis or septic shock	mNGS analysis of cell free DNA in blood plasma samples	Samples sequenced centrally using mNGS on an Illumina HiSeq2500; not performed in a near-patient context.
Optimization of sepsis therapy based on patient-specific digital precision diagnostics using next generation sequencing (DigiSep-Trial) [158]	Germany, 2020	randomised, controlled, interventional, open-label, multicentre	205 patients, two trial arms. Will characterise the effect of i) a combination of NGS-based diagnostics with standard-of-care microbiological analyses compared to ii) solely standard-of-care in patient with sepsis/septic shock, to see if inclusion of NGS improves outcomes.	mNGS analysis of cell free DNA in blood plasma samples	Samples are sequenced centrally using mNGS on an Illumina NextSeq; not performed in a near-patient context.
Optimising Treatment Outcomes for Children and Adults Through Rapid Genome Sequencing of Sepsis Pathogens (DIRECT) [159]	Australia, 2021	pilot prospective, non-randomised multicentre trial	Aims to demonstrate the feasibility and diagnostic accuracy of pathogen sequencing direct from clinical samples and estimate the impact of this approach on time to effective therapy when integrated with personalised software-guided antimicrobial dosing in children and adults on ICU with sepsis. Phase 1 will evaluate methods in 50 patients with blood culture-confirmed sepsis; phase 2 will apply methods to 50 patients with suspected sepsis	mNGS analysis of whole blood samples	Samples will be sequenced centrally using mNGS on Oxford Nanopore Technologies MinION; not performed in a near-patient context.

Table 9. Commercial NGS tests available or in development for sepsis pathogen identification

Test name	Summary	AMR genes detected?	Sample type	Sequencing approach	Turnaround times and suitability for near-patient testing	Development stage /regulatory approval
The Karius test (Karius, California, US) [160]	The Karius test can be used to detect pathogens causing both localised and bloodstream infections, including bacteria, fungi, viruses and other types of eukaryotes. It has been explored for use in the context of sepsis with studies showing that it can improve sensitivity of pathogen detection compared to culture. However the clinical benefits are still unclear, including the clinical contexts where it can add most utility in sepsis diagnostics [161].	No	Blood plasma	mNGS analysis	Test not performed in a near patient context, samples are shipped to central location. 24-hour turnaround times from sample receipt to result (not including shipping)	Commercially available for use in CLIA-certified/CAP-accredited laboratory
The iDTECT™ Dx Blood test (PathoQuest SAS, Paris, France) [162].	The iDTECT™ Dx Blood test is the first IVD CE-marked mNGS test for microorganism identification [162]. It can be used for diagnosis of bacterial and viral sepsis-infections, using sequencing before identifying pathogen sequences based on a curated database. It is also only currently recommended for use in immunosuppressed patients [163]. Similar to other mNGS tests, it has been shown to have a higher sensitivity but lower specificity than blood culture.	No	Whole blood	mNGS analysis	Test not performed in a near patient context, samples are shipped to central location. In a validation study time to result was 2-3 days, though it is unclear if this includes sample transport time [164].	Commercially available, CE IVD-marked

Test name	Summary	AMR genes detected?	Sample type	Sequencing approach	Turnaround times and suitability for near-patient testing	Development stage /regulatory approval
SepsiTest-UMS (Molzym, UK) [165]	Whilst based on sanger sequencing rather than NGS, Molzym's SepsiTest-UMD has been commercially available in Europe for several years for identification of bacterial and fungal pathogens in a variety of clinical specimens and for a variety of clinical indications, including sepsis [165]. The test was assessed by the National Institute of Health and Clinical Excellence in the UK in 2016 and reviewed in 2019, but although considered promising was not recommended for routine use in the NHS in England due to a lack of evidence on its utility, particularly in regard to clinical decision making [21]. Molzym also offer products to automate or semi automate sample preparation.	No	Whole blood	PCR amplification of 16S and 18S rRNA genes, followed by Sanger sequencing of positive results.	Test is suitable for performance in hospital laboratories with the required equipment e.g. PCR machine and sequencer. The PCR result is available after 4 hours of sample receipt. This is followed by sequencing; sequencing results may be available in 3 to 4 hours depending on the analyser used [21].	Commercially available in EU, CE IVD-marked. Not available in US but was granted FDA Breakthrough Device Designation in May 2021, for a variety of conditions, including sepsis
LiDia-Seq test (DNAe, London, UK) [75]	DNAe are developing a semiconductor sequencing platform (the same technology as Ion-Torrent sequencing) which automates the entire sequencing workflow from sample to result, so is suitable for use by non-experts. In theory, the platform will be able to detect 1200 bacterial and 90 fungal pathogens, with the company intending to use their own validated database.	Yes	Whole blood	16S and 18S rRNA sequencing, as well targeted sequencing of 30 common AMR genes	Specifically designed for near-patient settings, so tests can be performed at the point of need rather than in laboratories [75]. Intended to deliver results of stand-alone tests with 4 hours of sample receipt	In the prototype phase of development. Development is currently being funded by BARDA (U.S. Department of Health and Human Services), with a grant worth up to $51.9 million. Designated as an FDA breakthrough product in April 2020.

Test name	Summary	AMR genes detected?	Sample type	Sequencing approach	Turnaround times and suitability for near-patient testing	Development stage /regulatory approval
AutoSepT (Molzym, UK) [166].	Molzym recently announced they were developing a new platform called AutoSepT, in collaboration with Fraunhofer IZI-BB [166]. The test appears to be based on the same principles as the SepsiTest, but more user friendly and suitable for a wider-range of near-patient settings than the SepsiTest.	Not clear	Not clear	PCR amplification of 16S and 18S rRNA genes, followed by Sanger sequencing of positive results.	Designed to be fully automated and aims to deliver a turnaround time of sample to result in 4-5 hours.	In early phases of development.
Patho-Seq assay (Pathogenomix, US) [167]	Patho-Seq is designed for the rapid detection and identification of hundreds of clinically relevant bacteria for a broad list of clinical conditions and sample types, including sepsis. The Patho-Seq test consists of Pathogenomix's analysis platform which is applied to NGS data produced using wet-lab protocols developed for Patho-Seq. Any sequencing platform can be used.	No	Whole blood	16s rRNA sequencing	Designed so that the user produces the sequencing data, then uses the Patho-Seq platform so suitable for NPT in a hospital lab setting with sequencers. Turnaround time will vary depending on sequencing time. According to Pathogenomix data and analysis and reporting is performed in under 5 minutes.	A RUO platform currently exists but the clinical use Patho-Seq assay is not yet commercially available. US FDA granted Patho-Seq Breakthrough Device Designation in Jan 2022 for a variety of conditions, including sepsis.
Day Zero Diagnostics (US) [168]	Day Zero Diagnostics have produced several technologies for sample prep and computational analysis designed to be used alongside commercially available sequencing technology, as well as in-house databased for pathogen and AMR gene detection, to allow pathogen identification and AMR prediction direct from blood samples. One potential use indicated is for sepsis.	Yes	Whole blood	mNGS	Designed so that the user produces the sequencing data, then uses the Patho-Seq platform so suitable for NPT in a hospital lab setting with sequencers. Turnaround time not yet clear, aim is to achieve results within hours rather than days.	In development (product engineering phase)

6.1.5 Considerations for use of NGS tests for near patient sepsis diagnostics in LMICs

There are multiple factors to be considered when assessing if NGS based methods will be useful for pathogen identification in sepsis. Many of these apply to both HICs and LMICs, but some may be particularly important to consider from a LMIC perspective.

The clinical performance of a potential near patient NGS test

Data on the sensitivity and specificity of some NGS based tests for pathogen identification are being generated via research studies and will continue to be generated via clinical trials, such as those described in Table 8. However, these tests are being performed in a centralised context, therefore further validation will be required if tests are to be performed in near-patient settings. Importantly, the majority of research is currently taking place in HICs, where the profiles of patient populations with sepsis likely differ from those in LMICs. Further data are required on the prevalence of different pathogens, AMR genes and the burden of sepsis in each country or region in order to consider the implementation of NGS tests. Understanding the precise clinical scenario in which sepsis mNGS could have most utility, including the turnaround times and throughput required, will be necessary when evaluating the feasibility and suitability of this technology as a clinical test.

The clinical utility of performing the NGS test

Even though some NGS tests are commercially available for sepsis diagnostics there remains a lack of evidence on whether NGS testing improves sepsis outcomes for patients compared to current methods, especially in near-patient settings [21]. If near-patient NGS tests for sepsis are developed, it will be important to understand in which settings they deliver the most improved outcomes for patients i.e. in which clinical scenario (e.g. for adult or paediatric patients, or patients with particularly high risk of sepsis mortality) clinical pathway (e.g. as a standalone test vs following the use of other tests) and setting (e.g. in a well-resourced hospital laboratory vs a clinic with limited laboratory facilities and expertise). For example, an interviewee noted that even though NGS tests may deliver results faster than culture in experimental settings, the real-world timelines of test use may differ, depending on factors such as whether they can be completed within a single shift, or whether 24-hour laboratory services are available to allow the test to be carried out as soon as possible.

This applies to both HICs and LMICs, but it is possible that the areas of most utility differ for countries with different levels of resources. For example, healthcare facilities in low resource settings may lack alternative methods for pathogen diagnosis, such as culture. It is possible that use of NGS methods may deliver greater benefits to patients in these settings than in HICs, where culture is typically available.

However, in many situations similar clinical benefits may be obtained from implementing quicker and cheaper non-sequencing-based tests such as PCR and antigen tests, and sequencing could be seen as unnecessary. Similarly, the type of sequencing required will depend on the clinical need; mNGS data may be unnecessary for most diagnostic purposes, for which targeted methods would suffice. However, in situations where targeted methods or culture are not able to identify a causative pathogen, mNGS may be a useful second line diagnostic test. In some situations, there may be further utility to be gained from sequencing data if used beyond diagnostics, for pathogen surveillance and research, which could be taken into account when evaluating the utility of a particular test. However, the more comprehensive information produced, while useful for research or surveillance purposes, may also be in excess to that required for diagnosis. Simpler tests could be used to provide the same diagnostic information.

The overall sepsis diagnostic pathways in different LMICs also require consideration. Currently, pathogen identification is only determined following a diagnosis of likely sepsis. For many LMICs, there is limited information on how many patients are diagnosed and

treated for sepsis, but there is evidence that sepsis is often not recognised, for multiple reasons. Therefore, the utility of sequencing in improving outcomes may be limited, if sepsis cases are not identified in a timely manner in the first place.

Unless data are generated on the ability of sequencing to improve outcomes compared to current standard of care diagnostics for these different scenarios in LMICs, it will be more difficult to then perform the cost benefit analyses which will be required to justify implementing what is often considered a relatively expensive technology.

The feasibility of performing the NGS test in a specific setting

The facilities and resources available for near patient sepsis testing could vary from a well-equipped laboratory with sequencing equipment and expertise already available, for example, at a university hospital laboratory, to a rural hospital with less laboratory testing capacity and intermittent access to electricity and internet services. The different NGS approaches in development vary in design and complexity, some may only be suitable for established laboratories, others may be adapted for use in lower resource settings. For example, mNGS being developed using the ONT minion could also be adapted for other cheaper ONT devices such as the Flongle [169]. As research into the utility of NGS for sepsis diagnosis continues, and the need for near patient testing becomes clearer, then it will be important to take into account the healthcare facilities when designing sequencing-based tests.

6.1.6 Conclusions

Culture independent NGS testing has the potential to be used for pathogen identification in sepsis, providing a more rapid alternative to culture-based testing. This could be helpful in low resource settings where microbiological culture facilities required to identify them are often lacking. In addition, unlike current targeted molecular PCR panel tests available, pathogen agnostic NGS approaches do not require prior knowledge of the likely pathogens implicated in sepsis. This could make them useful in some LMIC settings where a broad range of pathogens have the potential to cause sepsis and/or the most common pathogens implicated in sepsis are unknown, making design and effective use of more limited pathogen specific tests challenging. mNGS tests may be particularly useful if a range of pathogen types are considered potential causes of sepsis, as all classes of pathogen (bacteria, fungi, viruses, and protozoa) can be sequenced.

However, currently, few NGS approaches for sepsis diagnosis in any context have been developed, let alone specifically for near patient testing, and evidence is still being generated on the validity and utility of different NGS approaches for pathogen identification. It will be necessary to ensure future research is designed to take into account the needs of LMICs rather than just HICs. NGS performed for other purposes apart from diagnostics could also be informative for helping address limitations in sepsis management, such as the lack of regional specific guidance for recognition and treatment. For example, more routine sequencing of septic cases in LMICs could be useful in identifying the most prevalent sepsis-causing pathogens in a particular region or healthcare facility. This in turn could inform the development of guidance for improved identification and treatment of sepsis cases and the creation of new diagnostic tests for sepsis-causing pathogens, such as PCR assays. Using NGS to assist sepsis-management in this way could improve sepsis outcomes in LMICS, at least in the short term, compared to using NGS as a diagnostic test [148].

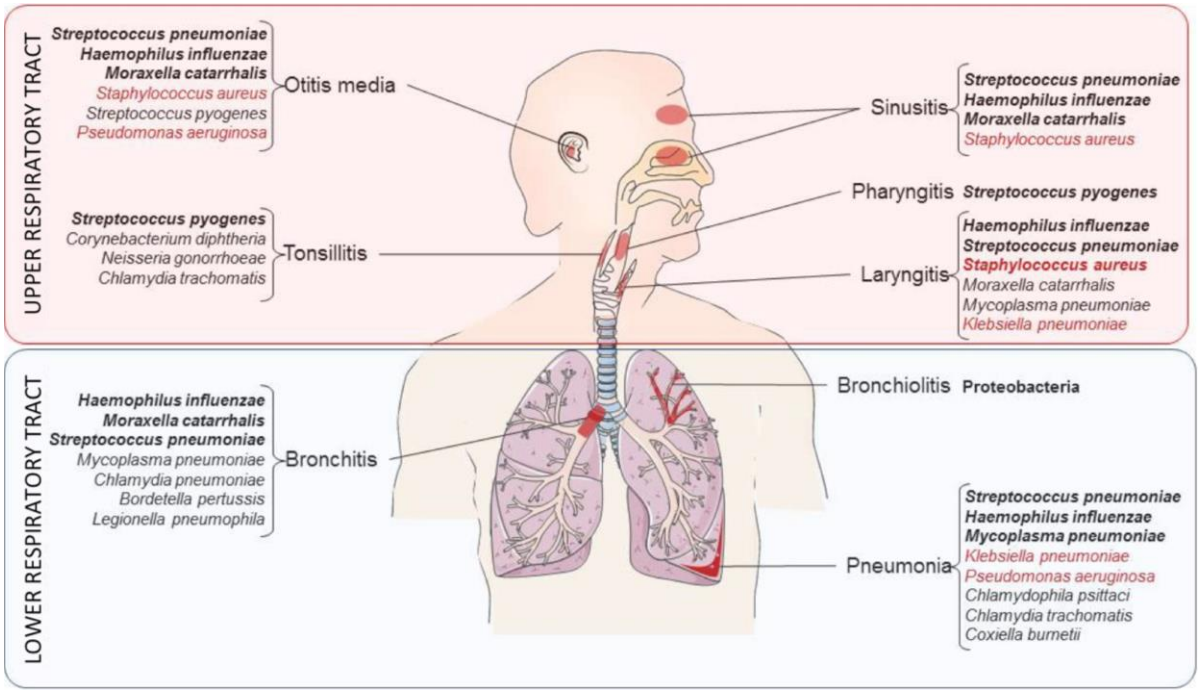
6.2 Respiratory diseases

Tuberculosis is a respiratory tract infection that has a strong evidence base for the use of sequencing in its management and AMR surveillance. In particular the Seq&Treat project that is currently underway is evaluating targeted NGS-based drug resistant TB diagnostic tests [79]. For this reason, respiratory tract infections other than tuberculosis are being discussed in this section. COVID-19 is also not discussed as there are other resources available, although it must be noted that co-infection with SAR-CoV-2 and other respiratory pathogens is occurring.

The human respiratory tract is divided into two spatial environments: the upper respiratory tract (URT), including tonsils, nasopharynx, oral cavity, oropharynx, and larynx; and the lower respiratory tract (LRT), including trachea, bronchi, and lungs. Acute respiratory infections cause approximately four million deaths per year globally [170]. A variety of viruses, bacteria and fungi inhabit the respiratory system and some can cause respiratory tract infections (Figure 6 shows bacteria that can be responsible for many of these diseases). Defining most of these patient diseases is difficult because the presentations connected with respiratory tract infections commonly overlap and their causes are similar. Most URT infections are of viral aetiology whilst the causative agents of LRT infections are viral or bacterial. Symptoms of a LRT infection range from shortness of breath, chest pain, difficulty breathing, coughing, fever, fatigue, wheezing to gasping for air and coughing up blood.

The majority of acute respiratory infection deaths and severe illnesses are due to LRT infections. Of the communicable diseases, LRT infections are the most deadly [171]. In 2016 they accounted for the death of nearly 2.4 million people of all ages worldwide [172]. LRT infections include community-acquired pneumonia, hospital-acquired pneumonia, bronchitis, bronchiolitis, and tracheitis. Malnutrition, air pollution as well as overuse of antibiotics have been identified as risk factors [172]. LRT infections are caused by a wide array of pathogens, such as bacteria, viruses, mycoplasma, and fungi, all of which present indistinguishable clinical presentations. The LRT can become a reservoir of bacterial pathogens that can lead to dramatic clinical outcomes [173, 174]. Furthermore, many bacterial pathogens involved in LRT infections are multi-resistant to antibiotic treatments and are considered as priority areas for antibiotic research and development by the WHO [175].

Figure 6. Bacteria responsible for the most common respiratory diseases. Summary of the bacteria responsible for respiratory infections in the URT (red zone) and the LRT (blue zone). Bacteria are grouped according to the respiratory niche they preferentially infect. Pathogens indicated in bold correspond to the main bacteria encountered in airways while the ones in red correspond to opportunistic pathogens. From [176].



Human respiratory viruses include a broad range of viruses that infect cells of the respiratory tract, that cause respiratory and other symptoms, and are transmitted mainly by respiratory secretions of infected people. Respiratory virus infections often cannot be differentiated clinically. These viruses differ in viral and genomic structure, the human populations susceptible to infection, disease severity, seasonality of circulation, transmissibility and modes of transmission. Common respiratory viral pathogen families include adenovirus, enterovirus, coronavirus, rhinovirus, influenza, and respiratory syncytial virus (RSV). Together, they contribute to substantial global morbidity [177], mortality [178] and concomitant economic losses [179] annually.

Invasive fungal infections are estimated to kill 1.5 million people per year worldwide [180]. Although they can affect any organ system, the respiratory tract is a prominent portal of access for filamentous fungi to enter the body and establish infection [180]. The airways are constantly exposed to environmental fungi. The risk of developing a fungal infection is increased in people with conditions such as asthma or cystic fibrosis, and in immunocompromised individuals. There is also the risk of developing a co-infection from a fungal infection when already suffering from another respiratory infection, for example pneumonia or COVID-19. Invasive fungal infections are an increasingly important cause of human morbidity and mortality with growing evidence that the lung fungal community (mycobiome) has a significant impact on the clinical outcome of chronic respiratory disease [181, 182]. As a consequence of culture-independent methods, especially NGS, several fungi that were previously undetected have been identified in human lungs [182].

6.2.1 Current approaches for respiratory infection diagnosis and treatment

In URT infections, such as otitis media and tonsillitis, a clinical examination is often sufficient to diagnose the disease and can lead to simple local microbiological sampling (e.g. swabs of the nose or throat). URT infections are generally mild in nature and most often caused by

viruses, sometimes with a bacterial component as in some cases of sinusitis and otitis media. In contrast, it is more complicated to detect and analyse LRT infections such as bacterial-related pneumonia based on etiologic examination and can require further examination such as x-rays, sputum and blood laboratory testing.

Since a variety of pathogens can cause acute respiratory infections, identification of the causative pathogen at an early stage of the disease can be important to implement effective antimicrobial therapy and infection control. Sample collection is needed for pathogen identification and appropriate administration of therapy to appropriately manage LRT infections. Initial sampling is non-invasive (and generally non-sterile) such as collection of phlegm that is coughed up. However, this technique can be more difficult in children as they have difficulties in providing these samples. An alternative is through inhalation of nebulised sterile saline solution. Urine collection can also be done and is easier to obtain when specific antigen tests for systemic infection are available [183]. Sterile, but often invasive, alternative collection techniques include: blood sampling; thoracentesis to remove fluid or air from around the lungs (in case of pleural effusion, a very common symptom of bacterial pneumonia); trans-thoracic needle aspiration; and bronchoalveolar lavages (performed with fibre-optic bronchoscope). For more than 40 years, a further method, protected specimen brush, has been considered as a reference test for diagnosis in pneumonia by limiting bacterial contamination from the URT as the brush is inserted into the respiratory tract [176]. The thin collection brush is protected by a sheath, when the brush reaches the desired area, it is extended to collect lung secretions and cells. After sample collection, specimen culture or PCRs can be done on the samples.

Without a definitive microbiological diagnosis, patients are initially treated with empirical broad-spectrum antibiotics. Guidelines recommend that such therapy should be refined or stopped after 2-3 days, once microbiology results become available. However, this is often not done if the patient is responding well or the laboratory has failed to identify a pathogen [184, 185]. Such extensive ‘blind’ use of broad-spectrum antibiotics is wasteful and constitutes poor antimicrobial stewardship. Up to 62% of community acquired pneumonia remain undiagnosed despite comprehensive diagnostic work-up [186]. Furthermore, in the absence of an identified microbial cause, clinicians may mistakenly classify the symptoms as a non-infectious inflammatory condition and in some cases prescribe empiric corticosteroids for treatment, which may result in worsening infection or reinfection [187].

Cell culture has been considered the “gold standard” for laboratory diagnosis of respiratory viral pathogens for decades. Culture is also the gold standard for detection of atypical (bacterial) respiratory pathogens, which is followed by identification and antimicrobial susceptibility testing by various manual or automated methods [188]. Culture-based diagnostics and susceptibility testing, in use for 70 years, have limitations as guides for the appropriate clinical management of acute infections, mainly because of their slow sample-to-result turnaround [189]. Clinical microbiology currently relies on culture and this is unlikely to change in the near future. Diagnostic fungal cultures are technically difficult and currently take 2–4 weeks of incubation [190], so there are considerable possible speed and accuracy benefits in switching to sequence-based methods of detection [191, 192]. Whilst best practice guidelines for blood culture in LMICs are available [151], these are not available for sputum culture. Automation of culture steps has been shown to be a key factor in reducing the time to obtaining a culture result [151] and such systems could be suitable for laboratories in low resource settings [193] .

Despite this, culture and susceptibility testing is slow, with typical turnaround times of 48–72 hours and low clinical sensitivity (low detection rate) [189]. Furthermore, the use of specialised culture media or anaerobic culture can stretch the capacity of routine clinical laboratories [194]. To maximise the impact on patient management, identification of clinically relevant antibiotic resistance genes as well as the infecting pathogen(s) is necessary.

Over the last few years, diagnostics for respiratory infections have evolved substantially, with the development of novel assays and the availability of updated tests for newer strains of pathogens. Newer laboratory methods such as nucleic acid tests and antigen tests are available for point of care [195], rapid diagnostics [176] as well as other laboratory tests [196]. These are highly sensitive and specific, and are gradually replacing the conventional gold standard tests, although the clinical utility of some of these assays is still under evaluation. Although culture-independent techniques, such as immunological assays and nucleic acid testing using PCR, are rapid and accurate, they require prior knowledge or assumptions regarding the types of pathogens.

6.2.2 Potential of near patient NGS as a diagnostic test for respiratory infections

As outlined in Section 4.1 there are a number of different sequencing approaches that can be utilised to investigate respiratory infections, each with their own benefits and limitations. They include pathogen agnostic sequencing (mNGS, targeted NGS and unbiased WGS of culture isolates) and pathogen specific (targeted NGS and targeted WGS). In theory any of these can be applied to respiratory infections if a sample of appropriate quality is available [194].

The ideal diagnostic test requirements discussed for sepsis (Section 6.1.2) are also applicable to respiratory infections, with further consideration for the samples being used and the complex microorganism environment of the respiratory tract complicating the analysis.

Pathogen specific targeted NGS is possible but requires prior knowledge or assumptions regarding the types of pathogen expected. It is difficult to differentiate different respiratory infections clinically as they can have similar presentations. Hence there has been most interest in pathogen agnostic sequencing techniques.

Most culture-independent microbiological techniques based on NGS studies have focused on the bacterial component of the microbiome, whereas other organisms such as viruses (virome) and fungi (mycobiome) have been less intensively investigated [182]. However, with genetic fingerprinting, it is now known that the diversity of the human mycobiome is greater than was expected because many species have not yet been cultured [182, 197].

Pathogen agnostic targeted sequencing can be achieved with the prior amplification of a gene such as 16S rDNA for bacteria and internal transcribed spacer (ITS) for fungi. For fungi, early studies focused on the 18S small subunit or the 28S large subunit rDNA, whereas recent studies prefer the ITS for its higher taxonomic discriminatory power [192]. The factors that determine the best sequencing target are within microbial communities are taxonomic resolution, coverage, accuracy, and amplicon length [182].

16S rRNA gene analysis of the whole bacterial community from biological specimens such as sputum and samples collected from throat swabs may provide direct information on the presence of pathogens without the need for culture [194]. 16S rRNA gene sequence analysis of DNA isolated from clinical specimens can better identify poorly described, rarely isolated or phenotypically aberrant strains. It could be routinely used for identification of mycobacteria (which are difficult to identify by culture-based methods) and can lead to the recognition of novel pathogens and bacteria that are resistant to standard clinical culture [198-200]. However, challenges arise from the accuracy of sequences in databases and may result in an uncertain definition of species when based only on sequence information instead of the traditional phenotypic classification of cultured microbial isolates.

Pathogen agnostic mNGS may serve as a new tool to overcome the shortcomings of conventional diagnostic methods that require culture or knowledge of which pathogens are present. The chief advantage of mNGS lies in its unbiased sampling and that there is no reliance on culture. It enables the simultaneous identification of all potentially infectious agents in samples and avoids defining the targets for diagnosis beforehand or relying on

them to grow in culture. It has the potential to overcome the shortcomings of both culture and PCR, by combining speed with comprehensive coverage of all microorganisms present. Compared to conventional tests, mNGS has a broader spectrum for pathogen detection in a single test, which streamlines clinical testing for pulmonary co-infection diagnosis. All of these reasons explain why there is a focus on the use of mNGS in respiratory infections.

Respiratory specimens present a difficult challenge for unbiased mNGS due to variable pathogen load, the presence of commensal respiratory tract flora, and the high ratio of host:pathogen nucleic acids present (up to 105:1 in sputum) [189]. In addition, the biological variation in sampling (timing of sampling, host DNA level, contamination, etc.) and the technical variation in methodology (different nucleic acid extraction methods, incomplete databases, differentiated bioinformatics tools, and unstandardised interpretation standards) limit its widespread use in a clinical setting.

In addition, it is recommended that every mNGS laboratory maintains a proprietary database that contains background microorganisms arising from normal flora or laboratory environments [201]. This could be particularly important with respiratory diseases as there is a background microbiome that is present in the respiratory tract. Most mNGS diagnosis platforms are based on short read sequencing and it can be challenging to determine whether detected antibiotic resistant genes originate from the genome of the causative agent rather than from the normal respiratory flora, or due to environmental contamination.

Depending on the sequencing technology, methods, and bioinformatics procedures, the turnaround time of mNGS is about six hours to one week from the time the sample is received, with an average of 48 hours [83].

6.2.3 Analytical performance of mNGS for respiratory infections

mNGS, as a culture-independent, unbiased, and hypothesis-free approach, has emerged as a possible diagnostic method for respiratory tract infections in recent years (Table 10) [187, 189, 202]. Current molecular tests for LRT infection diagnoses are usually pathogen-specific where a clinician selects relevant tests according to a patient's symptoms, which poses a challenge when novel or unexpected pathogens emerge. In contrast, mNGS can provide a comprehensive view of pathogens in a given sample, which enables the detection of novel and rare causative pathogens in the diagnosis of unexplained pneumonia [187]. For example, in early December 2019, severe unexplained pneumonia emerged in Wuhan, Hubei Province, China. On February 3, 2020, a novel coronavirus (SARS-CoV-2) was identified using RNA based mNGS that was determined to cause this pneumonia [203]. Compared to the time taken to identify SARS (five months), mNGS shortened the time taken considerably to five days for the accurate identification of the gene sequence of the virus [203].

A pathogen reference panel representing 30 different microorganisms of different types has been developed and assessed for performance in 17 sites across China [204]. This cross-workflow and cross-platform study is the largest effort to date to produce and analyse comprehensive reference datasets for metagenomics for pathogen detection in respiratory diseases. It will assist with setting up quality control and assessment procedures to enable mNGS to be implemented more widely.

Sampling methods for pulmonary / respiratory infections need further investigation to determine which is the most suitable for all NGS techniques and include: virtual bronchoscopic navigation, radial probe endobronchial ultrasound using transbronchial lung biopsy (TBLB), and bronchoalveolar lavage to collect specimens from patients with peripheral lung infections. However in mNGS, bronchoalveolar lavage fluid (BALF) has been found to detect a wider range of pathogenic microorganisms and can narrow the range of suspected pathogens. mNGS using TBLB sampling provides more accurate diagnostic results, while the hard-to-obtain specimens should be diagnosed using BALF sampling methods [187]. This could be due to a variety of reasons, such as the quality of the samples,

the community of pathogens present, or the abilities of the laboratory. Further research is needed to determine which types of samples and sampling methods are most suitable for different sequencing approaches. Pre-sequencing treatment such as filtration, centrifugation and nucleic acid removal using nucleases can enrich viruses and improve the sensitivity of viral metagenomics analysis in respiratory samples [205].

Table 10. The analytical performance of metagenomic (m) NGS in the diagnosis of respiratory tract infections. Abbreviations: AFS, acid-fast stain; BALF, broncho-alveolar lavage fluid; NA, no accessible; NPV, negative predict value; PPV, positive predict value; PSB, protected specimen brushes; TA, tracheal aspirate; ETA, endotracheal tube aspirate; MTB mycobacterium tuberculosis. From [187].

Study	Platform	Confirmatory tests	Samples	Sequencing	Sensitivity (%)	Specificity (%)	Concordance (%)	PPV (%)	NPV (%)
Integrating host response and unbiased microbe detection for lower respiratory tract infection diagnosis in critically ill adults [206]	Illumina HiSeq 4000	Clinical microbiologic testing	92 TA samples	RNA and DNA based mNGS	100	87.5	NA	NA	100 (In the validation cohort)
Nanopore mNGS enables rapid clinical diagnosis of bacterial lower respiratory infection [189]	MinION	qPCR	81 Respiratory samples (sputum, endotracheal secretions and ETAs)	DNA based mNGS	96.6	41.7	NA	NA	NA
mNGS for mixed pulmonary infection diagnosis [207]	NA	Conventional tests (smear, culture, pathology, GM test, Xpert MTB)	55 Pulmonary biopsy and BALFs	DNA based mNGS	97.2	63.2	39.9	83.3	92.3
The respiratory virome and exacerbations in patients with chronic obstructive pulmonary disease [208]	Illumina NextSeq 500	qPCR	88 Nasopharyngeal samples	RNA and DNA based mNGS	96	100	NA	82	100
Detection of Pulmonary Infectious Pathogens From Lung Biopsy Tissues by mNGS [209]	BGISEQ-500	Culture	20 Lung biopsy tissues	DNA based mNGS	Bacteria :100.0 Fungi: 57.1	Bacteria:76.5 Fungi :61.5	NA	Bacteria:42.9 Fungi:44.4	Bacteria:100 Fungi:72.7
mNGS versus Traditional Pathogen Detection in the Diagnosis of Peripheral Pulmonary Infectious Lesions [210]	BGISEQ-100	Culture, microscopic examination	240 Samples (lung tissue, BALF, and PSB)	DNA based mNGS	88.3	81.16	N	92.07	73.68
Clinical mNGS for diagnosis of pulmonary tuberculosis [211]	Illumina NextSeq CN500	Xpert, culture, and AFS	110 BALFs	DNA based mNGS	47.92	98.39	N	N	N
Retrospective Validation of a mNGS Protocol for Combined Detection of RNA and DNA Viruses Using Respiratory Samples from Pediatric Patients [212]	Illumina HiSeq 4000 and NextSeq 500	qPCR	19 Nasopharyngeal washings, 2 sputa, 2 BALF, 1 bronchial washing and 1 throat swab	RNA based mNGS	83	94	N	N	N
Application of mNGS for bronchoalveolar lavage diagnostics in critically ill patients [213]	BGISEQ-50	Culture	35 BALFs	DNA based mNGS	88.89	74.07	77.78	53.33	95.24
		Smear and PCR			77.78	70	73.68	70	77.78

At present, in the diagnosis of LRT infections, most studies focus on the comparison of the detection rates of all the pathogens (bacteria, virus and fungi together) detected by mNGS and traditional pathogen detection methods (Table 10). However, there are a few studies that compare the detection rates of each of the pathogen groups individually by mNGS versus traditional pathogen detection methods (Table 11) [205]. From these studies it is evident that clinicians still need to correlate clinical presentation to make accurate judgements when interpreting the reports, which can detect multiple pathogens. In some cases, viral infection may be limited to a specific area of a tissue or part of susceptible cells, so the detection of a single or very few viral sequences may also indicate viral infection. The appropriate value of the number of viral sequences remains to be determined [214].

Table 11. Detection rate of pathogens of pulmonary infection in mNGS versus traditional detection methods. From [205]. Traditional detection methods included bacterial and fungal smear and culture, Grocott's methenamine staining, acid-fast staining, and blood sampling to detect routine blood, inflammatory markers, specific antigen or antibody tests, PCR amplification or culture methods.

Pathogen	mNGS (%)	Traditional method (%)	Ratio (mNGS / traditional method)	Reference
Bacteria	42.1	17.5	2.4	[210]
	76.4	47.2	1.7	[215]
	75.0	13.8	5.4	[216]
	73.1	8.3	8.8	[217]
	85.7	42.8	2.0	[218]
	50.0	15.0	3.3	[209]
	65.0	20.0	3.3	[217]
Viruses	35.1	0.0	-	[210]
	53.8	41.0	1.3	[217]
	83.0	0.0	-	[216]
	88.3	0.0	-	[219]
	41.6	0.0	-	[215]
	84.3	28.1	3.0	[220]
Fungi	38.6	8.2	4.7	[210]
	93.0	19.8	4.7	[219]
	16.8	10.0	1.7	[216]
	90.4	4.7	19.2	[215]
	45.0	35.0	1.3	[209]
	71.4	4.8	14.9	[218]

Other than pathogens, most of the microorganisms in the respiratory tract have not been isolated or characterised in any detail [221]. The human respiratory tract is inhabited by niche-specific communities of bacteria, viruses and fungi. The ecological and environmental factors that direct the development of microbial communities in the respiratory tract and how these communities affect respiratory health are the focus of current research [221]. While mNGS is able to collect information on all microbes present, it may not always be clear which are pathogen(s) causing disease and which are commensal, meaning that sequencing information may not be informative for clinical care. As the evidence builds, it is possible that novel pathogen(s) associated with the presence of a disease will be identified. Currently, it is likely there will be a continued need for culture to provide direct clinical information on the presence of pathogens.

6.2.4 Current development status of NGS for respiratory disease

Clinical trials are currently ongoing to determine the clinical utility of NGS approaches for the diagnosis of respiratory infections, some of these trials exploring the use of sequencing of respiratory infections in adults and children are described in Table 12.

Table 12. Current clinical trials evaluating use of NGS for respiratory infections (excluding COVID and TB) from clinicaltrials.gov accessed 25 April 2022.

Trial name	Country, year registered	Trial type	Summary	Sequencing approach
Early Identification and Severity Prediction of Acute Respiratory Infectious Disease (ESAR) [222]	China, 2021	Interventional. Multicentre, prospective, and randomized study	Aims to determine the best way of early identification and severity prediction of acute respiratory infectious diseases. Patients with suspected acute respiratory infectious diseases will be enrolled into this study and receive two different diagnostic pathways.	mNGS
Screening Microorganism of Cryptogenic Mechanical Pneumonia Through Next Generation Sequencing to Lung Tissue Fluid (COPandNGS) [51]	China, 2020	Observational.	Aim to detect the pathogenic microorganisms in BALF and lung puncture fluid of cryptogenic organizing pneumonia (COP) patients to further clarify the correlation between the incidence of COPs and pathogenic microorganisms.	NGS
Pathogen Identification in Paediatric Hematopoietic Stem Cell Transplant Patients With Suspected Lower Respiratory Tract Infection [223]	USA, 2016	Observational. Multicentre cross-sectional	This study will correlate results with patient characteristics, clinical microbiology test results, and clinical outcomes in order to evaluate the utility of metagenomics NGS in improving the diagnosis of LRTI in our paediatric HCT population.	mNGS
Characterization of Respiratory Microbiota in Susceptibility to Viral Respiratory Infections (RESPIBIOTE) [224]	France, 2018	Interventional.	The aim is to determine the existence of respiratory microbiota profiles associated with the occurrence of viral respiratory infections influencing the clinical expression of virus and to determine the role of the respiratory microbiota in the occurrence of bacterial superinfection which will justify an early antibiotic treatment.	mNGS targeting the V3-V4 hypervariable regions of the 16S RNA gene. Using a MiSeq Illumina.
mNGS -Guided Antimicrobial Treatment in Early Severe Community-Acquired Pneumonia Among Immunocompromised Patients (MATESHIP) [225]	China, 2022	Interventional.	This study aims to determine whether mNGS (using LRT specimen) guided antimicrobial treatment improves clinical prognosis of severe community-acquired pneumonia patients with immunocompromised conditions when compared with conventional antimicrobial treatment.	mNGS

The INHALE Randomised Controlled Trial (RCT) is currently underway in the UK [226]. As part of the early stages of the trial three rapid diagnostic systems were evaluated on respiratory specimens from ICU patients, namely: the PCR-based BioFire FilmArray, Curetis Unyvero platforms and rapid sequencing with the Oxford Nanopore Technologies MinION [227]. Based on the results, the FilmArray Pneumonia Panel (the 'FilmArray test') was selected as the best

performing test to carry forward into the RCT. The overall trial aim is to show that clinical and safety outcomes for patients whose treatment is guided by the FilmArray test molecular diagnostic are non-inferior compared to standard care, but that altered prescribing leads to improved antimicrobial stewardship.

Current and future point-of-care and near patient respiratory viral infection tests exists, they are mainly nucleic acid amplification tests and none are sequencing specific [195, 228]. However, rapid genome sequencing analysis accommodates the fast development of reliable in-house and commercially available nucleic acid amplification tests reagents shortly after an outbreak onset. Some NGS based tests for respiratory infections are available, however these are research use only and are not commonly used, even in higher income countries (examples in Table 13).

Table 13. Examples of next generation sequencing based tests for respiratory infections (research use only).

Name	Summary	Sequencing method
Illumina's Respiratory Virus Oligo Panel [229]	Targets ~40 respiratory viruses This method allows for near-complete sequence data of targets and opens up applications such as variant analysis for viral evolution or viral surveillance.	Target enrichment through hybrid–capture methods for Illumina instruments
Illumina's Respiratory Pathogen ID/AMR Enrichment Panel Kit (Powered by IDbyDNA Explify) [76]	Targets >280 respiratory pathogens and 1200+ antibiotic resistance alleles Universal Detection of Respiratory Pathogens — Delivers comprehensive detection of critical pathogens, including SARS-CoV-2. The respiratory pathogen panel simultaneously detects RNA and DNA for 180+ bacteria, 50+ fungi, and 40+ viruses. Detection of Antimicrobial Resistance Genes — Accurate prediction of resistance of 16 common respiratory pathogens to 60 relevant antimicrobials based on detection of > 1200 associated AMR markers Strain Typing of Critical Pathogens — Report full genome coverage of SARS-CoV-2 and Influenza A/B viruses to surveillance of new variants and lineages	Target Enrichment, Target Enrichment, Targeted DNA Sequencing, Targeted RNA Sequencing for Illumina instruments
CleanPlex® Respiratory Virus Research Panel V2 [230]	For the detection, research, and surveillance of COVID-19 SARS-CoV-2 virus, Influenza (flu A/B), and respiratory syncytial virus (RSV) (A/B), enabling complete genome sequencing of the SARS-CoV-2 virus and concurrent influenza and RSV sub-typing with highly sensitive detection.	Targeted sequencing, amplicon-based on Illumina sequencing platforms
Celemics' Comprehensive Respiratory Virus Panel (CRVP) Respiratory Virus Targeted Sequencing Panels [231, 232]	Testing of 39 strains for 9 different virus types; human adenovirus, bocavirus, human rhinovirus, coronavirus, human enterovirus, influenza (flu a/b), parainfluenza virus, and respiratory syncytial virus (RSV).	WGS on Illumina platforms
TwistBiosciences Respiratory Virus Research Panel [233]	Targeted against the reference sequences for 29 common human respiratory viruses, including: Coronavirus (CoV), Influenza virus, Adenovirus, Bocavirus (hBoV), Enterovirus, Metapneumovirus, Parainfluenza (hPIV), Human rhinovirus (HRV), Measles morbillivirus (MeV), Mumps virus (MuV), Rubella virus, Respiratory syncytial virus (RSV). Additional probes were designed to incorporate diversity from 77 additional rhinovirus strains, and to target diverse genomes representing each major influenza A and B outbreak since 2000.	Targeted enrichment

6.2.5 Role of sequencing in surveillance and diagnostic tests of respiratory diseases

Apart from a role in the diagnosis of respiratory infections NGS approaches can play an important role in surveillance and public health measures. In addition, being able to use NGS to determine which pathogens and variants are circulating within a community can provide valuable information for the development of diagnostics. A case study on the use of metagenomics on clinical respiratory samples from a Kenyan hospital is outlined below. The provision of sequencing within various hospitals and communities for surveillance purposes can be invaluable in monitoring which pathogens are present, causing disease, AMR and relevance of local diagnostic tests.

Title: Identification of missed viruses by metagenomics sequencing of clinical respiratory samples from Kenya (2022) [234]

Aim of study: Evaluate the ability of NGS to identify viruses missed by the diagnostic panel by evaluating samples from patients with respiratory disease which tested negative for common respiratory viral pathogens.

Project Summary: Investigated clinical respiratory samples from a single location in Kenya that had failed to return a diagnosis with the local PCR diagnostic panel. Demonstrated the utility of direct deep sequencing of clinical respiratory samples to identify virus genomes circulating in a resource-limited country.

Sequencing requirements: Total nucleic acid extraction and dsDNA conversion were performed. The method includes centrifugation and DNase treatment to remove free non-encapsidated DNA, reverse transcription with non-ribosomal random hexamers avoiding rRNA targets followed by sequencing on Illumina HiSeq 2500, generating 2–3 million 250 nucleotide paired-end reads/sample. Standard Illumina libraries were prepared for each sample. (Further described in [235])

Key findings: The approach revealed three categories of missed diagnosis.

1. The virus was in the diagnostic panel, but locally circulating strains differed at diagnostic primer/probe sites. This is remedied by updating the diagnostic panel with locally appropriate primers
2. The virus was detected but below the cut-off for a positive diagnosis. The current cut-off for positive diagnosis is PCR Ct < 35.0; yet the assays did not show Ct-values of negative diagnoses, hence limiting the interpretation of findings. Further studies are needed to determine assay sensitivities to update positive/negative cutoffs
3. The virus was not in the diagnostic test panel. Due to practical constraints, it is not feasible to include all potential pathogens in a diagnostic panel, however findings from NGS analyses may support the decision to modify diagnostic panels accordingly. A limitation of this NGS approach is the threshold below which a virus does not yield identifiable sequences. Although 30.5% of the samples returned a viral diagnosis, 69.5% failed to yield classifiable viral sequences. It is expected that future improvements in NGS methods will increase the fraction of new diagnoses allowed by these methods.

Next steps: Although viral NGS would be expensive to apply for all cases, the data from this study provide evidence for applying agnostic viral NGS to improve local diagnostics and the frequency that this would be suitable. Sequencing information can inform which genetic targets should be included in new panels. It is expected that declining diagnostic sensitivity occurs over time due to virus evolution with altered primer target sequences, movement into the region of undetectable variants or viruses not on the diagnostic panel. Each round of NGS would result in a revised diagnostic panel, adjusted for local sequence variation and new viruses.

6.2.6 Conclusions

The use of NGS could impact on better selection of the therapies for the management of respiratory diseases, reduce the use of broad-spectrum antimicrobials, potentially avoiding invasive procedures and assist in surveillance processes. Whilst still in the research stages of development and not yet ready for deployment, there is some evidence of the potential utility of near patient pathogen sequencing tests. Many severe respiratory infections generally require admission of patients to a hospital or clinic and therefore this may be the optimal site to deliver such tests in the future.

NGS can identify novel or unexpected pathogens, identify fastidious pathogens as well as co-infections. Currently, mNGS has been increasingly employed for the unbiased detection in clinical samples from infectious patients. However, few NGS approaches specifically for near patient testing in respiratory infections have been developed, and evidence is still being generated on the validity and utility of these approaches for pathogen identification. Sequencing can reduce the time-to-result from days to hours and with the increasing portability of sequencing devices, it could be done closer to the patient, reducing the time spent sending samples to a central laboratory. NGS tests beyond syndromic panels currently deployed in diagnostic laboratories is possible. Nonetheless, many challenges remain to be addressed before metagenomic sequencing can be widely adopted, especially the current lack of evidence that this approach can improve patient care.

6.3 Malaria

Malaria is caused by eukaryotic *Plasmodium* parasites of which *P. falciparum* and *P. vivax* are the two most concerning species. The parasite is transmitted by *Anopheles* mosquitoes, with different stages of the lifecycle occurring in mosquitoes and humans. Following infection, parasites develop in the human liver before proliferating in the blood, where gametocytes are formed that can infect mosquitoes during a subsequent bite; it is also at this stage that symptoms and diagnosis can occur. Some malaria species, including *P. vivax* can enter a dormant stage in the liver and later reactivate, causing symptoms and infectious disease long after initial infection [236].

Malaria is a challenge particularly in sub-Saharan Africa, but also affects countries in South-East Asia, Eastern Mediterranean, Western Pacific, and the Americas. *P. falciparum* is dominant in Africa and causes the most severe disease while *P. vivax* is dominant in countries outside of sub-Saharan Africa. Sub-Saharan Africa has the most cases and deaths from malaria; in 2020 there were 627,000 estimated deaths from malaria globally, with 96% of these being in African countries and 80% of those being children under 5 years old [237]. Children and pregnant women are among those most at risk of severe disease and mortality [238].

6.3.1 Malaria diagnosis and treatment

Malaria diagnosis is typically achieved either by rapid diagnostic tests (RDTs) that detect antigens or microscopy of blood smears, which is considered the gold standard. RDTs are cheap and do not require expertise or infrastructure so are the predominant diagnostic test in lower resource settings, while microscopy-based diagnosis is used where facilities and expertise are available. No molecular diagnostic tests have been recommended for malaria diagnosis by the WHO to date, although PCR-based diagnosis is sometimes used in research settings [239, 240].

Despite being the most used diagnostic for malaria, there are some challenges associated with RDTs. Parasites have mutated to evade detection by RDTs, resulting in false negative results. RDTs that detect *P. falciparum* histidine-rich protein 2 (HRP2) are widely used but *pfhrp2/3* gene deletions mean parasites are able to evade detection [241]. In addition, a study found that, of *P. falciparum* infections detected by ultrasensitive PCR, only 28% were detected by microscopy and 34% were detected by RDTs, highlighting a potential need for improvement when trying to eradicate disease even in asymptomatic individuals with low parasite burden, although the clinical relevance of this is currently unknown [242].

The best treatment for malaria currently is artemisinin-based combination therapy (ACT), which is recommended by WHO for the treatment of *P. falciparum*. This leads to the elimination of blood-stage parasites but does not impact dormant parasites in the liver. Treatment must be given within 24 hours of symptom onset to prevent serious potentially life-threatening malaria. For severe malaria artesunate treatment is recommended, followed by a standard course of ACT. Combination therapy is used to reduce the emergence of resistance, with the use of monotherapy being discouraged by WHO unless in exceptional circumstances [243].

6.3.2 Malaria control and elimination

Malaria control and elimination programmes have proven successful in some countries, with 11 countries being certified as malaria-free by the WHO in the last 20 years and an estimated 1.7 billion cases being averted over 20 years due to control programmes [240]. The most successful methods for controlling malaria have been the use of insecticide-treated nets, indoor residual spraying, prophylactic drugs, and antimalarial treatments. However, in order to achieve and maintain malaria-free status a number of additional approaches are required, which includes surveillance to track the disease and enable targeted public health approaches [237]. The 2021 WHO global technical strategy for malaria identified increased surveillance as being a fundamental component of national malaria strategies. The strategy has set goals to reduce malaria cases and mortality by at least 90% between 2015 and 2030 and to eliminate malaria from at least 35

countries during the same timeframe, as well as preventing re-establishment in malaria-free countries [244].

While there are a number of effective antimalarial drugs available, drug resistance is a recurring challenge. Resistance has already been described to drugs used in ACT in several areas, particularly in the Greater Mekong Subregion. Resistance often develops in this region where there is relatively low prevalence of malaria but high antimalarial usage, before spreading to nearby areas. There are concerns ACT resistance may spread to Africa, although this has not been detected to date [3]. The ability to track drug resistance is a key requirement of malaria control programmes, to ensure targeted responses where resistance is emerging and to tailor treatment in areas with resistance.

In 2021, the RTS,S/AS01 vaccine was the first malaria vaccine to be recommended by the WHO, which is for the prevention of *P. falciparum* malaria in children living in regions with moderate to high transmission. This recommendation came following a pilot in Ghana, Kenya and Malawi, which found the vaccine to be safe and cost effective in these settings. Additionally, deployment of the vaccine did not reduce the use of other malaria control measures, such as the use of bed nets, or delay the seeking of healthcare when experiencing a febrile illness. However, it is likely the parasite will evolve to evade the immune response developed via vaccination in the same way it has developed drug resistance. There are further malaria vaccines in development, including two in late stage clinical trials, and WHO has created preferred product characteristics and guidelines to inform vaccine development [240].

6.3.3 Sequencing the malaria parasite

Sequencing approaches for the parasite, mosquitoes, and humans are all being explored. Mosquito sequencing is used to track insecticide resistance and human sequencing is used for pharmacogenomics to identify those who will have an adverse reaction to anti-malaria drugs and for investigating susceptibility and resistance to severe disease. Parasite sequencing allows for the tracking of resistance variants and for detailed investigation of transmission patterns, as well as parasite species identification. Within a near patient setting, parasite sequencing likely has the most utility initially due to the smaller genome size (23Mb for *P. falciparum*, 230-284Mb for *Anopheles* mosquitoes [245, 246]) and the useability of parasite sequencing results to inform public health decision making, including altering treatment and testing approaches.

Currently, parasite sequencing is generally carried out on dried blood spots with amplification to increase parasite DNA and reduce the proportion of human DNA in the sample. The use of dried blood spots from finger-pricks makes this approach amenable to sample collection in the field, as it is simple to collect and store, and refrigeration is not required. Previously, amplification of the parasite genome was a challenge but newer methods, such as the selective whole genome amplification method developed by MalariaGEN in their amplicon sequencing toolkit, have meant that sequencing can be carried out from dried blood spots without the requirement for larger volumes of blood for human DNA depletion protocols, which require more expertise and infrastructure [247, 248]. For other pathogens, culture may be used for amplification but this is complex and time consuming for malaria parasites and some species are not possible to culture [249].

The *P. falciparum* genome consists of 14 chromosomes and, even though the genome is much smaller than that of the mosquito vectors, it is still larger than bacterial or viral pathogens, meaning WGS can be costly and challenging. For this reason, alternative approaches, such as panel sequencing of specific, pre-defined genes of interest, are likely to be preferable, with most studies currently being based on amplicon sequencing [249].

Although sequencing has the potential to add value, particularly for malaria surveillance, it is not routinely carried out in many countries with malaria transmission. Increasing sequencing surveillance may help to achieve global aims to control and eliminate malaria.

6.3.4 Malaria sequencing considerations

Sequencing can be more complex than other diagnostic and surveillance approaches and there are several challenges associated with implementation. These include considerations around whether the data produced by sequencing is sufficient to justify the resources required; which sequencing technology and approach is most appropriate; how to handle data generation, storage, analysis, and sharing; and how to ensure standardisation and reporting of the findings [169, 249]. Many of these considerations are universal for the implementation of sequencing in lower resource settings and are discussed in Chapter 3. Considerations vary depending on the specific requirements and resources of different countries and sites, meaning optimum approaches may vary. However, this may complicate standardisation and hinder multinational malaria transmission investigations [169].

Standardisation is a key consideration for malaria sequencing, both within and outside of countries, to ensure data can be pooled and compared to enable tracking of transmission and spread of drug resistance across and between countries. Reporting processes also need to be in place to ensure results are given to national decision makers and malaria control programmes in an easily understandable format to inform public health measures in near real-time [169].

Sequencing can give detailed information on parasite lineage and drug resistance but there may be simpler and cheaper approaches that could be used in settings where this information is not required. Considerations about the optimal approach to give sufficient data without increasing resources unnecessarily are required, with cost effectiveness studies potentially able to highlight where sequencing may add value [169]. This also applies to which sequencing approach is taken, for example WGS is associated with increased cost, time, and requirement for expertise, whilst the relatively simpler and less resource-intensive targeted panel sequencing approach may be sufficient (Chapter 3) [249].

In addition to the above implementation challenges, the *P. falciparum* genome has a high AT content of 81% and contains many areas that are difficult to sequence, including extended repetitive regions [250]. This can be particularly challenging when aligning short read sequencing, with long read sequencing providing easier and more accurate alignment[63]. In addition, some drug resistance genes are long and complex, such as the k13-propellor domain associated with artemisinin resistance; long read sequencing can cover the entire gene while short read can give unreliable coverage and be challenging to align when multiple reads are required to cover the whole gene. For this reason, long read sequencing technologies, such as ONT, might be preferable, but these are often associated with higher error rates than short read technologies, such as Illumina sequencing (Chapter 3). Detection of alternative alleles on Nanopore sequencers may be less reliable than on Illumina sequencers when sequenced at a low depth, therefore there needs to be careful consideration of which protocol to use [63].

6.3.5 Sequencing use cases

A number of areas where sequencing of the malaria parasite could add value have been identified. Some of these use cases overlap, with a single sequencing dataset able to address several purposes, depending on the context.

Sequencing for diagnosis and treatment decisions

Sequencing for malaria diagnosis is not being utilised currently due to the availability of cheap RDTs and microscopy. Indeed, very little research is being done in this area, with more focus on sequencing for malaria surveillance, where there is potential to add value compared to other approaches. Even in high resource settings, such as the UK malaria reference laboratory that is responsible for diagnosis and surveillance of imported cases, diagnosis is carried out using microscopy, with PCR-based molecular tests not being routinely carried out [251].

Expert interviewees thought use of sequencing in malaria diagnostics would be an option in the future if the resources and time required to carry it out were reduced. It was considered that sequencing cannot compete with RDTs due to their speed and low cost. Instead, RDTs that can detect drug resistance as well as diagnose malaria were seen as being the optimum diagnostic solution in the future. The tailoring of individual treatment based on resistance profiles may also help to reduce the spread of drug resistance in the future. However, this would rely on having alternative antimalarials available if resistance was detected, which is often not the case currently.

Sequencing can help to inform the development of other diagnostic tools. For example, CRISPR-based diagnostics are being developed, based on knowledge gained from sequencing data, that are able to identify asymptomatic cases with low parasitaemia, distinguish between malaria species, and potentially detect drug resistance in a lateral flow assay, although these are currently still at a research stage [252, 253].

There have also been studies investigating the reuse of samples that were previously used in RDTs as a sequencing input for molecular surveillance. This would help to ensure comprehensive sample collection, even in remote regions, and so increase coverage without increasing workload for local healthcare workers. Proof of concept studies found that it is possible to carry out amplicon sequencing on these samples and gain useful insights for malaria surveillance. The use of RDT samples was less reliable than using dried blood spots and the sample and library preparation was more expensive and time consuming, due to the requirement for more complex PCR protocols with longer running times and the large number of samples that did not give usable data. However, the potential to easily collect a large number of samples and run these in a high throughput manner means that this approach may still prove beneficial and cost-effective [254, 255]

Sequencing for drug resistance surveillance

Tracking the spread of drug resistance is a key use case for targeted sequencing in the management of malaria, with only a few genes requiring sequencing meaning this can be done using a small panel at lower cost and with less complex data analysis and storage requirements compared with WGS or more extensive panels. Currently drug resistance tracking is conducted using expensive and time intensive treatment efficacy surveys, with differing levels of coverage and success between regions. Because of this, sequencing may enable better coverage for surveillance [169].

Sequencing of known drug resistance genes can help to inform public health decisions around which antimalarials to use within a region and help to focus resources to areas with high levels of drug resistance to reduce malaria cases and contain spread. Prioritising elimination programmes to areas with high levels of drug resistance may help to delay or prevent the spread of drug resistance to endemic areas with high case numbers and mortality, including sub-Saharan Africa [169, 249, 256].

However, there are challenges linking genotype and phenotype, meaning the relevance of new mutations may not be clear and will require characterisation, particularly in species other than *P. falciparum* which is the best studied [169]. There may be potential for interplay between sequencing and treatment efficacy studies, whereby sequencing of potential drug resistance genes is used to identify areas of priority for efficacy studies, which can in turn help to validate the potential resistance genes.

Sequencing for diagnostic test failure

Detecting the presence of *pfhrp2/3* gene deletions, which result in false negative HRP2 RDTs, is required to ensure all malaria cases are identified and can be treated. These deletions can be found at high prevalence in some areas; the Horn of Africa is an area of particular concern with over 50% of cases being missed by these RDTs. The WHO has recommended switching to an alternative diagnostic method, such as microscopy or alternative RDTs, in areas where ≥5% of

parasites harbour *pfhrp2/3* deletions [241]. However, molecular surveillance is needed to detect these gene deletions to identify areas where alternative RDTs should be used.

Alternative RDTs are limited but some are recommended by WHO, including those that detect *Plasmodium* lactate dehydrogenase (*p*LDH), which can be universal or species-specific, and aldolase, which detects all species. *p*LDH RDTs have been found to be slightly more specific but less sensitive than HRP2 RDTs for *P. falciparum* and may also be less stable at high temperatures [243, 257].

Sequencing for vaccine development and surveillance

Sequencing has the potential to aid vaccination efforts in two ways: firstly, by identifying parasite genes that are stable across the population, which might make their encoded antigens candidates for vaccine targets; and secondly, by monitoring vaccine targets in areas where vaccination programmes are in place to identify if resistance develops. Currently, only a single vaccine is recommended by WHO but, if further vaccines are recommended in the future, sequencing of parasite genes encoding antigen targets may help to select which vaccines will be most effective in any particular region [240, 249].

Sequencing for transmission and outbreak surveillance

Tracking infections and transmission patterns is an area where sequencing can provide additional information compared to traditional use of incidence, prevalence, and travel history investigation. This may be particularly useful in areas with low prevalence which are working towards elimination to determine if cases are being transmitted locally or imported from elsewhere.

Identifying imported cases typically relies on travel surveys and/or mobile phone data, which can be sporadic and unreliable. In addition, parasite transmission is not solely based on human movement, with mosquito vectors also transmitting across borders. Genetic data provides information on the relatedness of parasites, without having to estimate human or mosquito movement to trace infection, so can be used to more accurately track transmission, which in turn can inform public health responses [169].

Distinguishing recurrent infections from reinfection is another area of interest. *P. vivax* parasites can enter a dormant phase in the liver before reactivating and sequencing can help identify whether disease is caused by a new infection or by reactivation by looking at parasite lineages. This again helps to inform public health response and control programmes, particularly aiding in assessing the uptake and efficacy of treatments that are able to eliminate dormant parasites in the liver, such as primaquine. Knowledge of the prevalence of recurrent infections, which may account for up to 80% of *P. vivax* infections in some areas, and efficacy of these treatments, which are associated with significant side effects in some people, can help to inform optimum treatment strategy at a population level [169, 258].

Sequencing also allows for the accurate identification of malaria species, including zoonotic malaria species, such as *P. knowlesi*, that may be misidentified as other species by other approaches. Knowledge of which species is causing an infection can aid in malaria surveillance and control programmes and can help to inform treatment decisions, as different treatments are recommended for different species [259].

However, due to the nature of the malaria parasite lifecycle, whereby sexual reproduction occurs within mosquitoes and asexual proliferation occurs in humans in the blood stream, and the potential for co-infection with multiple parasite strains in both humans and mosquitoes, particularly in areas with high transmission, determining relatedness between parasites from genomics can be more challenging than for viral or bacterial infections. Models need to be developed that can reliably and accurately deconvolute these complex interactions [249].

6.3.6 Sequencing in the care of malaria – examples

MalariaGEN is a network of malaria researchers that are building shared malaria genomic data resources to enable access to sequencing protocols and to encourage data sharing. This includes parasite, mosquito, and human sequencing data. The network includes over 200 people, including clinicians, researchers and those involved in malaria control programmes, across over 40 endemic countries. Parasite sequencing data is being collected from 30 countries across Africa, Asia, and South America [260-262]. Members undertake sequencing, predominantly amplicon for parasites and WGS for vectors, and make data available for research, including into vaccine design and drug development, and as a reference for future sequencing, with the broad aim of helping achieve malaria control and elimination. MalariaGEN also encourages collaboration between partners, improving understanding of malaria globally and increasing funding opportunities, while working with national malaria control programmes (NMCPs) to inform public health decision making [11].

Currently, all sequencing of vectors and parasites for MalariaGEN is undertaken at the Wellcome Sanger Institute in the UK; however, the network is working towards setting up in-country sequencing, with regional hubs that can support surrounding areas and neighbouring countries with less infrastructure. Protocols used at the Sanger Institute are being adapted for lower resource settings and they are providing support, particularly on data analysis pipeline development and bioinformatics training.

GenRe-Mekong is a surveillance project that is a part of MalariaGEN and aims to produce data that can inform public health decision making in the Greater Mekong Subregion by producing data on linages and drug resistance and giving this to NMCPs and other stakeholders in an easily accessible format, termed a genetic report card. The project is taking place in Bangladesh, Cambodia, Laos, Myanmar, Thailand, Vietnam, and India, with sequencing being undertaken at the Sanger Institute [263]. The GenRe-Mekong project was based on the SpotMalaria framework, which is a genetic surveillance platform that provides methods for genotyping drug resistance genes and genetic barcodes for molecular epidemiology. The genetic barcodes are based on 101 SNPs with high geographical variation that allow for basic transmission investigations. Although older protocols used mass spectrometry and capillary sequencing, updated protocols use amplicon sequencing on Illumina MiSeq sequencers, with a prior selective whole genome amplification step to amplify the parasite genome and reduce the proportion of human DNA [3].

Results from GenRe-Mekong, alongside confirmatory *in vivo* treatment failure data, led to changing of national drug policy in Vietnam, with several provinces altering frontline treatment due to the identification of drug resistance. However, these approaches are only possible when genetic surveillance is undertaken continually and with sufficient coverage across the whole region. Additionally, WGS is required to identify new drug resistance variants to then be included in the gene panel, although this can be done at less frequent intervals and with less population coverage than the standard genetic barcoding approach. Challenges in getting sufficient DNA from dried blood spots were identified and samples with low level parasite load, below the level of detection by microscopy, were not possible to genotype in most cases [3]. This may mean that patients with low parasitaemia are missed and, although this is unlikely to be of clinical relevance, this may impact on surveillance into transmission patterns, particularly in areas with low transmission. However, WHO only recommends the use of nucleic acid amplification tests in research settings because of their ability to detect cases with low parasite load, stating that more research is required into the role of these low level infections on public health before they can be recommended more widely [14].

There are no available commercial NGS-based tests for malaria. Surveillance is being undertaken in research settings, as there is no evidence supporting the clinical use of NGS-based tests. Expert interviewees stated that the area that may have the most potential in the future for commercialisation could be an NGS panel or solution that can rapidly and simultaneously test for drug, RDT, and vaccine resistance, allowing for real-time tracking of these by NMCPs to help inform control measures and treatments.

6.3.7 Near patient sequencing of malaria parasites

Often surveillance of drug resistant parasites is sporadic and focused on areas that already have facilities to carry this out, rather than in, areas where it would have the most impact [264]. Near patient sequencing approaches may help to overcome this challenge by reaching more remote and smaller populations that are not as accessible for centralised testing approaches. NPPS for malaria would likely be based on panel testing, due to the large and complex genome of malaria parasites. A mobile NPPS approach could move between priority areas and outbreaks as they emerge to enable informed decision making within a short timeframe or fixed local laboratories could be utilised for routine surveillance.

Regular near patient malaria panel sequencing would help to detect the emergence and track the spread of drug resistance and RDT failure, allowing for more targeted interventions and rapid responses to minimise the spread of resistant parasites, alongside surveillance of vaccine failure when these become more widely available. NPPS of malaria parasites for transmission surveillance may have most utility in areas with low levels of transmission that are striving for malaria elimination to enable detailed transmission investigation and to distinguish local transmission from imported infections, although these would be more difficult to implement as they require more complex pipelines that have not yet been developed to a sufficient standard. For example, modelling approaches are needed that are able to take into account interactions between a number of factors, including the potential for coinfections and reactivation, and transmission dynamics.

The Mobile Malaria project, which aims to highlight malaria research being conducted in Africa and to teach mobile sequencing to local partners in low resource settings, can be seen as a proof of concept that shows that near patient malaria sequencing is possible, even though the aims of the project were not focussed on malaria control [265]. In this project they showed that mobile Nanopore sequencing could be used to detect drug resistance. This has further been confirmed by two studies that showed Nanopore MinION sequencers can be used to sequence malaria parasites both for drug resistance monitoring and parasite surveillance, albeit in centralised laboratory settings rather than in a near patient setting [63, 266].

The Mobile Malaria project also highlighted some other NPPS considerations. It provided workshops teaching local scientists how to run the sequencing and laboratory protocols and by using a local data analysis pipeline involving a MinIT and MinION sequencer and a laptop, showed that rapid sequencing results may be possible without internet access. However, it was highlighted in interviews that a mobile NPPS approach may not be optimum for routine NPPS for malaria care and that setting up smaller local laboratories would be preferable, except perhaps in cases where there is a large outbreak in a remote region that is not easily covered by a local laboratory.

Another study identified seven theoretical use cases for sequencing that would provide benefit for NMCPs, including ideal time to results. They identified that there is utility in having rapid turnaround times of less than a week, namely for the detection of resistance on an individual level and the identification of imported cases in malaria-free or low transmission areas. Turn around times of less than a month could have utility for investigating drug resistance gene spread, transmission intensity to prioritise interventions, identifying outbreaks, and investigating what is driving local transmission [256]. NPPS may therefore be a useful approach in these use cases to ensure results are available within a meaningful timeframe to allow for rapid response.

Malaria NPPS has not yet been implemented routinely and key research questions remain to be addressed, these include:

- Development of models for transmission based on genetic data that take into account complex transmission dynamics, including the potential for reinfection, reactivation, and coinfection
- Development of optimised protocols based on Oxford Nanopore Technologies sequencing, which may allow for more rapid implementation of NPPS where there is insufficient

infrastructure for Illumina sequencers. This includes investigating the impact of the increased error rate on results. Most protocols developed to date, including those developed by MalariaGEN, have been for Illumina sequencers
- Investigation into optimum sample type, for example dried blood spots or samples from RDTs, in a near patient context. This may vary depending on setting and resource, with extraction from RDTs requiring less work for healthcare staff but more resources for sample and library preparation
- Continued investigation of the genes involved in drug resistance, including linking phenotype and genotype
- Routine genomic surveillance of the parasite using WGS to identify new genes of interest associated with drug or test failures
- More evidence to show where sequencing data for malaria can add value to local and national programmes particularly for public health purposes.

6.4 Differences in test development for various diseases

While there is an interest in and development of sequencing-based tests for a range of diseases and conditions, the stage of development of these tests for specific diseases varies for reasons that are not always clear. In sepsis, sequencing is being actively explored in clinical trials and via commercial test development for rapid pathogen identification, while in respiratory disease and malaria sequencing is being used more in the research setting, or tests are in development but have not had clinical approval.

The reasons for the differences in the stage of sequencing test development between these disease areas could be:

Different priorities. Sepsis has unmet clinical need to terms of the need for rapid diagnosis and managing the challenges of delays to this process caused by culture. While respiratory diseases are recognised as an area where improved diagnostics are required, the focus globally has been on TB due to the disease incidence worldwide and the burden of AMR. Sepsis and respiratory disease are also prevalent in HICs, which can act as a driver of innovation, compared with malaria, which primarily affects LMICs only. The priorities for sequencing in malaria are as a surveillance tool to track resistance and disease transmission, as cheap and effective rapid diagnostic tests are already available.

Cost and utility of sequencing compared to cost and utility of existing tests. Sequencing might be too costly for what is required, particularly where there are cheaper alternatives available. For example, many respiratory diseases are caused by viral pathogens for which there are reliable PCR tests available. In malaria, there is currently not an unmet clinical need in terms of diagnostics, which can be done using cheap RDTs.

Sampling requirements. The ease of sampling can also support test development, for example blood and upper respiratory tract samples are easier to collect compared to lower respiratory tract samples, which could act as an enabler or barrier to test development

Causative pathogen and requirements for culture. The requirement for culture is particularly an issue for sepsis, where bacterial pathogens represent a high proportion of causative organisms. In lower resource settings, access to and availability of culture facilities can be a challenge.

Looking ahead, pan pathogen tests developed for a particular purpose – e.g. sepsis diagnostics – could be modified or repurposed for a different disease area, e.g. respiratory disease. Sepsis could be considered as a useful example or test case for how test development in other diseases could be pursued.

7 Cross-cutting themes and conclusions

Near patient pathogen sequencing (NPPS) facilities can offer more time sensitive and localised sequencing in a range of scenarios, for example to support diagnosis and subsequent clinical decision making, inform real-time genomic epidemiology and public health management of disease outbreaks, and support ongoing disease surveillance. NPPS can take place in decentralised yet static laboratories in hospitals or smaller health facilities such as clinics. Alternatively, NPPS can be carried out using mobile or field laboratories, operating in potentially extreme conditions.

While NPPS facilitates more rapid time to results, there is no agreement amongst experts or in the literature on a standard for the time frame within which a sequencing approach would automatically be considered near patient. Accurate and rapid diagnosis – as discussed for sepsis – would require sequencing facilities to be located very close to the patient to minimise sample transport time and other delays not related to the sequencing process. However, in many other circumstances, near patient sequencing is delivering an improvement in terms of access to testing and time to results compared to more remote facilities or are delivering results in locations where sequencing facilities were previously inaccessible or unavailable.

In some clinical or public health scenarios, there may be a clear need for near patient sequencing, in others a more centralised sequencing model may be sufficient. However, in countries where there is not an established network of centralised sequencing laboratories, or these are under-resourced or challenging to access, NPPS facilities – particularly mobile or field laboratories – may help to bridge the gap and facilitate use of sequencing in such settings.

In some circumstances, the term ‘near patient’ could be misleading, such as when mobile laboratories designed to bring sequencing facilities closer to the point of need may be located some distance away from the point at which patient samples are collected. For example, expert interviewees stated that during the 2018 Ebola epidemic in the Democratic Republic of Congo, mobile laboratories deployed to Goma, in the East of the country, were located hundreds of kilometres away from the villages where patient samples were collected. However, Goma was still geographically much closer to these villages than the capital city of Kinshasa located over 2,500 kilometres away in the west of the country, where the main sequencing facilities were located. Therefore, sample transport times and subsequent time to result was much quicker than if samples had to be transported to Kinshasa. In situations such as these, where sequencing facilities are sparse and there are limited options in terms of delivering sequencing to a large geographical area, the terms nearer-patient, point of need or even decentralised sequencing may more accurately capture the situation.

There are particular circumstances in which NPPS may be a priority for implementation, for example in AMR surveillance through the use of WGS on cultured samples as outlined by the *Global Antimicrobial Resistance and Use Surveillance System* (GLASS). Laboratories planning to include WGS in their services should have adequate infrastructure for isolating organisms from clinical samples, culturing them and extracting DNA from cultured isolates – currently, clinical microbiology relies on culture and this is unlikely to change in the near future [25]. Those developing sequencing in near patient contexts should also consider whether and how available culture facilities can be improved or introduced if they do not exist. Automation is one improvement in culture methodology that could be suitable for laboratories in low resource settings and has been shown to be a key factor in reducing the time to obtaining a result [151, 193].

Genomics has also been recognised as a global priority for surveillance, as outlined in the 2022 WHO report *Global genomic surveillance strategy for pathogens with pandemic and epidemic potential* [267]. This document outlines a global strategy for integrating genomic surveillance into public health globally, building on the advances made during the COVID-19 pandemic. The five high level objectives outlined in the report to support the strategy are:

1. Improve access to tools for better geographic representation
2. Strengthen the workforce to deliver at speed, scale and quality
3. Enhance data sharing and utility for streamlined local to global public health decision-making and action
4. Maximise connectivity for timely value-add in the broader surveillance architecture
5. Maintain a readiness posture for emergencies.

While the achievement of these objectives will have a greater impact on the use of sequencing more broadly, some of the NPPS-related issues discussed in this report are particularly relevant to objectives 1 and 5.

Objective 1 has a focus on appropriate and sustainable technology and infrastructure adapted to country needs. While highly centralised laboratories can be an effective model in HICs that also have good associated infrastructure in terms of sample transport and data management, more flexible models are required in LMICs and/or lower resource settings. The examples outlined in Chapter 5, mobile laboratories, demonstrate how such adaptable infrastructure could be achieved. The EAC mobile laboratories, for example, have already been used for surveillance for a number of diseases, including COVID-19.

Objective 5, maintaining a readiness posture for emergencies, raises considerations around how to optimise sequencing capacity. The flexible models of delivering NPPS, as outlined in this report, could have a role to play. For example, mobile or seasonal/semi-permanent sequencing capacity could be deployed close to the sources of potential outbreaks or as part of surveillance strategies for pathogens where incidence has a seasonal pattern. In order to ensure that overall sequencing capacity is not underutilised in anticipation of emergencies, careful consideration will be required as to what sequencing services could be repurposed to meet emergency needs. This may require joint projects and further collaboration with other sectors, for example One Health partners, universities and wider academia, and commercial companies.

7.1 Desirable characteristics of near patient sequencing technologies

In theory, all sequencing approaches (Chapter 4) could be performed in a near patient context. In practice, approaches that are designed for direct from sample analysis without need for prior culture may be more suitable, to reduce the additional infrastructure and expertise needed and reduce turnaround times. Other factors such as ease of use, accessibility, affordability, complexity of data analysis and sequencing need (e.g. diagnostics, surveillance) will also influence the type of approach used. For example, diagnostics may require lower throughput sequencing devices but require sequencing approaches with high accuracy and faster turnaround times, whilst for surveillance the turn-around time may be less critical but a higher throughput may be required. However, these needs are not unique to NPPS, and instead determined by the requirements of the specific application for which sequencing is being used.

Sequencing equipment, precise method choice, and other workflow design factors including workflows specifically intended to simplify more complex approaches, will also have an impact on the overall choice of the sequencing-based assay.

Many of the common features required of sequencing platforms to carry out NPPS in lower resource settings are similar to the requirements for sequencing in centralised yet resource limited laboratories. These are:

- Ability to operate in situations with unreliable electricity and temperature control
- Ability to operate with an unreliable cold chain
- High degree of automation in terms of workflow processes and procedures
- Access to offline analysis tools
- Low maintenance equipment with limited moving parts – this includes sequencers and accessory laboratory equipment
- Workflows that limit the amount of accessory equipment needed

- Equipment that can be moved or transported without requiring extensive set up and recalibration
- Minimal reliance on other laboratory functions, e.g. culture
- Flexibility in terms of throughput

7.2 Technical characteristics of current sequencing platforms

Illumina and Oxford Nanopore Technologies are the principal platforms that have been used for NPPS. ONT technologies are particularly amenable to field use and have been used in the majority of NPPS examples outlined. Smaller Illumina machines can be used in more static situations but are less amenable to transport and field use.

Main technical characteristics of ONT platforms (MinIon, GridION):

- Rapid and flexible sequencing with sequencing in real-time
- Relatively mobile with demonstrable use in a range of environments and settings
- Relatively accessible for users with less experience and expertise in sequencing
- Smaller sequencers are relatively inexpensive and comparatively low cost
- More limited barcoding for multiplexing of samples resulting in a higher cost per sample, however at lower throughput overall costs are comparatively lower
- Long read sequencing, although higher base-call and overall error rate than short-read technologies.

Examples outlined in this report show that ONT technologies can meet the needs for NPPS in some circumstances, although these still have limitations, for example in situations where accuracy is an important consideration. ONT's MinION sequencer has been used most extensively for near-patient and field sequencing, as demonstrated through available examples in the literature and feedback from expert interviewees. Improvements in long read sequencing quality and growing interest in the use of these platforms for diagnostics have raised a number of considerations for the evaluation of these technologies. Areas for future consideration include assessment of bioinformatics tools, benchmarking datasets, models for scalability of data processing, error correction, and adaptation for RNA or methylation sequencing data [42].

Main technical characteristics of Illumina platforms (iSeq, also MiniSeq and MiSeq):

- Most widely used sequencing platforms with extensive expertise and established workflows
- Range of machines with different specifications to meet different purposes
- Requirements for environmental control of temperature, humidity and vibrations
- More expensive overall, but lower cost per sample at higher throughputs
- Longer sample preparation and sequencing run times
- Short read sequencing with limited accuracy in complex genomic regions.

Experts reported that there are challenges around the use of Illumina platforms in lower resource or field settings, for example due to the availability of consumables and servicing or repairs for machines, the suitability of workflows, and the robustness of equipment in more challenging environmental conditions.

Another difference between the platforms is that some will have an instrument computer with a user interface versus the requirement for a high powered laptop or PC that is connected to the sequencer. Having an integrated computer allows some within-instrument data storage and analysis. Alternatively, data can be transferred to a laptop, the cloud or a PC for further analysis. Some ONT sequencers require high powered laptops or PCs with user interface software installed to be connected to the sequencer to power, initiate and monitor a sequencing run. Considerations for the computation requirements, particularly if restricted to a particular manufacturer, could increase costs.

7.3 Technological developments needed to support NPPS

Our research and feedback from experts highlighted the following technological areas where research and development are required:

- Reliable and cost-effective automation of wet laboratory processes. This can include tools developed for sample preparation, DNA extraction and library preparation (e.g. ONT VolTRAX)
- Streamlining and simplification of workflows and standardised and/or validated sequencing protocols/workflows. This could be achieved through reducing the number of stages required, flexibility around the number of samples that can be processed at once, minimal additional equipment required to carry out the workflow, and clear quality control stages. One 'sample to result' solution currently at the prototype phase is the LiDia-SEQ device, being developed by DNAe, which is planned to automate the entire sequencing workflow within a device designed for use by non-experts in non-laboratory settings [75]
- Reagents with a long shelf life which can be stored at ambient temperatures
- Sequencing instruments that can function smoothly after being moved, in response to vibration or where there is an unreliable or unstable electricity supply
- Any specifically designed sequencing methods and reagents, such as isothermal amplification (e.g. LAMP) that remove the need for thermocycling equipment for DNA amplification. Methods have been adapted for certain diagnostics (e.g., ONT LamPORE for SARS-CoV-2) [94]
- Tools to facilitate data analysis, use, management and sharing. This will include a combination of offline vs. cloud-based resources that offer flexibility in situations where internet access is sporadic, unreliable or non-existent
- Reduction in the volume of consumables required to carry out the workflow and limiting the volume of waste (particularly plastic) generated. Innovative products are being developed such as ArrayTape, a polymer strip embossed with reaction wells for sequencing. A 384-well array in Array tape is manufactured using approximately 1/7th of the plastic needed for a 384-well microplate [268]
- Further development and validation of culture-free sequencing approaches for specific applications and pathogens, especially when intended for diagnostic use, can help facilitate sequencing in settings where microbial culture is not available. Culture-free approaches can enable the faster turnaround times often desired for near-patient applications, and pathogen-agnostic approaches have many advantages in their own right. However, though culture-free approaches show promise, they are still challenging to develop, due to often very-low levels of pathogen nucleic acids present in clinical samples [269]. mNGS methods may provide the most comprehensive sequencing data but are further complicated by high levels of host DNA in samples. mNGS results are also particularly influenced by sample type and processing, microbe type and environmental contamination [270]. Development of procedures for standardisation and quality control is required, for example through the use of standardised reference panels [204]. All potential tests developed will also require validating in the populations in which they are intended to be used since it cannot be assumed that tests that appear to work well in HICs will work as intended in LMICs. Further data gathering is also needed on interpretation of mNGS methods, and accurate identification of pathogenic organisms.

7.4 Factors affecting implementation of near patient pathogen sequencing

A number of factors that affect the implementation of NPPS were identified, many of which impact on the implementation of sequencing more broadly, but which can be particularly acute in lower resource and/or remote near patient settings.

7.4.1 Availability of equipment and consumables

Procurement challenges, including the relative costs of equipment and consumables, and delays in procuring them, are a well known and ongoing problem for many LMICs. Expert interviewees reported that the time taken to receive consumables and equipment, and the costs of these relative to those paid by HICs, are a major challenge. The availability of open-source software and protocol adaptation to reduce consumables, such as those used by the OpenLab container laboratories (Chapter 5) will help.

Estimates of the cost of consumables from general laboratory equipment (e.g. pipettes, centrifuges), plastics (e.g. tips, tubes), reagents and computational requirements (e.g. laptops, software) are variable and will depend on the specific setting of the laboratory and the jurisdiction in which it resides. For example, availability of consumables – for example plastic tubes and pipettes – was affected during the COVID-19 pandemic when laboratory supplies were in high demand, with some countries introducing mandates on their export. This has an impact on countries that lack in-country manufacturing, or do not have the purchasing power to mitigate shortages.

Consumable costs are often higher in LMICs because of higher costs of shipping, customs formalities and taxes, and profit margins for local companies and distributors because manufacturing is not carried out within country. Many sequencing companies do not have a presence in a number of LMICs. There is also a need to accommodate fluctuating exchange rates when establishing costs. The cost of appropriate cool or cold chain shipping and storage needs to be considered. This may significantly affect the ability to maintain the quality of sequencing reagents; reagents also have a specific shelf life, sometimes hindering stockpiling.

Ongoing sustainable use of sequencing technologies in mobile settings requires an effective procurement and ordering process by which laboratory staff can order and receive consumables and reagents (including PPE) in a timely and effective manner. To help manage this process, procurement could be supported by experts with regional understanding of the challenges.

Many methods are protected by patents or licensing by a single proprietary company limiting access to enzymes and reagents underpinning a technology or method. Most sequencing platforms rely on kits developed by the sequencing company and it may not be possible to find or develop lower cost alternatives. These kits can be categorised as closed or open systems – closed systems require reagents from a specific producer, typically ensuring high quality reagents but creating dependence on a single supplier, which can restrict access to a particular technology. This also relates to methods in the sequencing workflow, such as isothermal amplification, limiting innovation in-country to develop sequencing solutions to meet their needs.

7.4.2 Availability of expertise

A range of personnel are required to establish and operate sequencing facilities and expert interviewees identified the poor availability of bioinformatics expertise in lower resource settings as being a particular challenge. Some sequencing approaches are more technical in nature than others, for example assembling a genome may require specific expertise as opposed to using a simpler test that sequences specific genome targets to confirm whether or not a pathogen is present. Workflows that are more complex will require a higher level of expertise from clinical scientists and laboratory technicians. Experts identified the need to simplify workflows and increase automation – which result in shorter training requirements – as one way to help mitigate this challenge by enabling a wider range of people to be trained. Instrument manufacturers can support this by continuing with their research efforts to develop easier to use equipment and training tools.

Using a ‘train the trainer’ model to increase expertise in lower resource settings was seen by experts as being the most sustainable approach. However, developing this can take time, particularly in the case of complex bioinformatics, meaning sustained investment in training is

required. Expert interviewees identified different levels of expertise and experience required by individuals prior to training to perform sequencing. Limited molecular biology expertise was highlighted as a major barrier to this training.

The Chan Zuckerberg (CZ) Biohub in the US is an example of an organisation helping to establish genomic capabilities in LMICs, especially metagenomics and use of sequencing for surveillance [271]. They provide training and support not just limited to the establishment of sequencing facilities but ongoing as needed throughout projects, including access to bioinformatics databases and support, e.g. the free CZ ID cloud based platform for analysis of metagenomic data. Whilst these projects are not specific to near-patient sequencing, by providing countries with the ability to perform sequencing to meet regional needs, this may facilitate near-patient sequencing as and when it becomes feasible and useful to do so.

Access to engineers to service sequencing equipment was also highlighted as being a challenge in LMICs, meaning that equipment can remain unusable for long periods before being fixed. There are a number of ways this challenge could be addressed, including developing in-country technical expertise and training laboratory technicians in basic equipment maintenance. Support from sequencing companies in such initiatives is likely to be essential.

Another relevant consideration is that research carried out in higher resource settings might not be relevant or applicable in lower resource settings. Therefore, protocols and expertise developed in these settings may not be immediately applicable in lower resource settings and adaptation will likely be required. Appropriate training that takes these adaptations into account and allows for collaborative working to support efforts to optimise pipelines in settings with different requirements.

7.4.3 Infrastructure and logistics

The number of possible NPPS technical approaches means that there will be wide variation in the infrastructural resources required in terms of the electricity and water supply, computational power, internet access, transport of equipment and samples, and laboratory footprint. Some approaches will also rely more on additional laboratory facilities such as pathogen culture and PCR, which will limit their use to more established settings as opposed to mobile laboratory or field use. There is a need for evidence gathering around the use of sequencing to replace established microbiology methods, for example phenotypic antimicrobial resistance testing, and it is likely that both approaches will still be required while this evidence base is developed for a range of pathogens. Mobile laboratories may be deployed for more than one purpose, for example, the mobile laboratory deployed during the Ebola epidemic in North Kivu, DRC for sequencing was also used for PCR diagnostics and its limited size resulted in PCR taking precedence, limiting the resource available for sequencing [56].

A further important consideration is that despite feedback from experts stating that culture-free sequencing methods will be important in the future, culture is still a requirement for the accurate diagnosis and resistance profiling for many pathogens and can also be used as a sample for sequencing. The WHO GLASS report on WGS for surveillance of antimicrobial resistance states that laboratories planning to use WGS should have adequate infrastructure for isolating organisms from clinical samples, culturing them and extracting DNA from cultured isolates [25]. If suitable culture and sequencing facilities are available, this sequencing can be carried out in a near patient setting. Some countries may choose to prioritise developing culture facilities in line with their sequencing facilities to support these efforts, given the utility of culture in helping to diagnose infections which could be caused by a number of pathogens e.g. respiratory diseases or febrile illnesses. Innovation in culture techniques, for example efforts to automate processes and reduce costs, will have a positive impact [193].

7.4.4 Data analysis and management

There are many considerations around the capture of NPPS data and how it is analysed, stored, reported and shared. In many cases these are the same as the wider considerations of pathogen data management which have been discussed extensively elsewhere [272-274]. There are, however, a number of the implementation issues specific for NPPS.

When NPPS approaches are delivered in resource limited environments, there will be a need for the most appropriate tools to analyse and store data. In terms of data analysis, proprietary tools such as BaseSpace/DRAGEN (Illumina) or EPI2ME (ONT) could prove useful in circumstances where bioinformatics expertise to develop your own analysis tools is limited. Consideration could be given to community developed tools, whereby local experts are trained to develop bioinformatics pipelines specific to the contexts and purpose of use. However, these types of approaches require extensive investment in terms of training the relevant experts and supporting development of future experts in the mid- to long-term.

One key consideration for NPPS is that many bioinformatic tools are cloud-based, which limits or extends the time needed to perform analyses where internet connection is unreliable. Offline data analysis options are available. For example, one study used GENEIOUS, proprietary sequence analysis software with a user-friendly interface, to develop an offline local BLAST search without the requirement for bioinformatics expertise [36]. Alternatively, software packages are available for bioinformatics with the sequencing equipment and sequencing platforms may come with these integrated into the device (i.e. Illumina iSeq 100, MinION Mk1C).

The uploading of data to cloud-based services and/or shared via international databases may create concerns around data governance and as with all sequencing data, international efforts are underway to manage the challenges around data sharing and ensuring that those who generated the data benefit from its use by others.

7.5 Conclusions

Near patient pathogen sequencing has utility in genomic surveillance during disease outbreaks and could in the future inform the development of more rapid diagnostics of disease or conditions such as respiratory infections.

In some cases, NPPS is being deployed via mobile laboratories in the absence of established centralised sequencing facilities, or in situations where sequencing facilities are too geographically distant to support timely genomic epidemiology. There is, therefore, a question around how extensive NPPS facilities would need to be, should more reliable and centralised ‘hub’ sequencing laboratories be established in countries where they are currently few in number.

Many of the features of sequencing platforms that make them suitable for use in a near patient setting also benefit delivery of sequencing in lower resource settings more broadly – for example automation, lower consumable costs, and open-source analysis software. Near patient sequencing approaches can therefore provide useful examples of what can be achieved, such as innovation in mobile and field laboratories which demonstrates that sequencing can be delivered under extreme conditions, with limited available resources.

The use of NPPS in mobile and field laboratories also raises considerations around longer-term capacity building. These laboratories have, in some cases, bridged the gap between emergency and more established services. There are considerations around how near-patient approaches such as these could be embedded as more established services, for example to help manage seasonal disease outbreaks. New services could be established from scratch, or mobile laboratories could augment existing centralised laboratory services and be dispatched where needed. During the 2013-16 Ebola epidemic, and during the current (as of May 2022) COVID-19 pandemic, there has been much capacity building for NPPS facilities and efforts will be needed to ensure that this is not lost. The opportunity to repurpose these resources to cover endemic

diseases, surveillance and local public health priorities should be considered. These could include AMR and OneHealth surveillance. While efforts in this area are underway, establishing services such as these will require reliable long-term funding to ensure stability and ongoing service provision.

The role and value of NPPS has been demonstrated to a certain degree but considerable further development is needed to make it a routine option for health authorities in low resource settings. This development will require close and effective collaboration between companies, researchers, global health agencies and most importantly the healthcare professionals and public health authorities who will be commissioning and using these facilities in the future.

8 Appendix

8.1 Abbreviations

A/C – alternating current
ACT – artemisinin-based combination therapy
AFS – acid-fast stain
AMR – antimicrobial resistance
AST – Antimicrobial Susceptibility Testing
BALF – bronchoalveolar lavage fluid
BOMB – Bio-On-Magnetic-Bead
BSL – biosafety laboratory
BGI/MGI – Beijing Genomics Institute and MGI Tech
cDNA – complementary DNA
CMOS – complementary metal-oxide semiconductor
CRISPR – clustered regularly interspaced short palindromic repeats
DNA – deoxyribonucleic acid
EAC – East African Community
EAC-ML – EAC mobile laboratory
EBOV – *Zaire ebolavirus*
EMLab – European Mobile Lab
ETA – endotracheal tube aspirate
EU – European Union
FHIR – Fast Healthcare Interoperability Resources
GCLP – Good Clinical Laboratory Practice
GLASS – WHO Global Antimicrobial Resistance Surveillance System
GOARN – Global Outbreak Alert and Response Network
HIC – high Income Country
HRP2 – histidine-rich protein 2
HTS – high-throughput sequencing
ICU – intensive care unit
IVD – *in vitro* diagnostics
INRB – Institut National de Recherche Biomédicale
ITS – internal transcribed spacer
LAMP – loop-mediated isothermal amplification
LIMS – laboratory information management systems
LMIC – low and middle income countries
LRT – lower respiratory tract
MALDI-TOF MS – Matrix-assisted laser desorption/ionisation-Time of Flight mass spectrometry
mNGS – metagenomic next generation sequencing
MTB – *Mycobacterium tuberculosis*
NGS – next generation sequencing
NMCP – national malaria control programme
NPPS – near patient pathogen sequencing
NPT – near patient testing
NPV – negative predictive value

ONT – Oxford Nanopore Technologies
OT2 – Opentrons liquid handler
PC – personal computer
PCR – polymerase chain reaction
PGM – Personal Genome Machine
*p*LDH – *Plasmodium* lactate dehydrogenase
POC – point of care
POCT – point of care testing
PPV – positive predictive value
PSB – protected specimen brushes
qPCR – quantitative polymerase chain reaction
QT-NASBA – quantitative nucleic acid sequence-based amplification
RCT – Randomised Controlled Trial
rDNA – recombinant DNA
RDT – rapid diagnostic test
RNA – ribonucleic acid
rRNA – ribosomal RNA
RRML – rapid response mobile laboratories
RSV – respiratory syncytial virus
rt-qPCR – quantitative reverse transcription polymerase chain reaction
SARS – severe acute respiratory syndrome
SARS-CoV-2 – severe acute respiratory syndrome coronavirus 2
SNP – single nucleotide polymorphism
TA – tracheal aspirate
TB – tuberculosis
TBLB – transbronchial lung biopsy
UK – United Kingdom
UPS – Uninterrupted power supply
URT – upper respiratory tract
V – volt
W – watt
WGS – whole genome sequencing
WHO – World Health Organization

8.2 Overview of advantages and limitations of NPPS technologies

	Oxford Nanopore Technologies	Illumina	Thermo Fisher Ion Torrent
Advantages	Rapid and flexible - particularly useful for sequencing smaller genomes. Sequences read in real time – allowing for termination when user determines enough reads have been generated Smaller sequencing units can be purchased at low cost Mobile sequencing – the small size and high portability of some systems means that these can be used in the field Some reagents use dry lyophilisation technology and do not require cold storage meaning they can be safely stored for a set period in environments where refrigeration is not possible or unreliable A simple user interface and analysis platforms for base calling and analysis are under development Direct RNA sequencing and also methylation sequencing are possible Automated sample preparation technology available. Relatively inexpensive at low throughput	Comparatively low-cost sequencing at high throughput, where many samples are being processed High raw read accuracy and read depth generating high accuracy data One of the more commonly used systems for high resolution genomic analysis allowing collaborative development of expertise and advances. Many genetic or research laboratories already possess these systems, and bioinformatics pipelines are relatively well-established High levels of sample multiplexing are possible, meaning a high number of samples can be run at once Option of targeted and WGS approaches	Relatively inexpensive at low throughput Comparatively short sequencing runs enable faster return of results Low substitution error rate Some systems facilitate a highly automated workflow for easy adoption and consistent application of sequencing Option of targeted and WGS approaches Longer individual reads

	Oxford Nanopore Technologies	Illumina	Thermo Fisher Ion Torrent
Limitations	More limited barcoding of samples limits multiplexing for some applications with higher costs per sample for high-throughput sequencing applications Some techniques use reagents requiring cold storage meaning these approaches can only be used with reliable cold storage and reagents may have a short shelf-life (<3 months). Dry lyophilised kits require intact storage bags in order to maintain low humidity Raw signal output files are very large which has an impact on data storage, data availability for analysis and subsequent data deposition into databases Bioinformatics expertise often needed to develop analysis pipelines, although examples of use by non-experts once available Lower read coverage than short-read platforms Higher base-call error rate and overall error rate Current automation limited in terms of sample barcoding and added cost of sequencing	Longer sequencing run time Most platforms are large and costly to purchase, some require specialised infrastructure for safe use Short reads limit accuracy in complex genomic regions and opportunity for identification of the genomic context of mobile genomic elements Installation of Illumina machines will have a number of environmental considerations, including vibrations, placement requirements, and control of temperature, humidity, elevation and air quality.	Lower throughput in comparison to other NGS technologies, therefore comparatively expensive at high throughput. Environmental constraints relating to humidity, temperature, vibration and air quality. Higher sequencing error rate Higher cost per sample Shorter reads than are possible with other NGS technologies able to perform paired-end sequencing Fewer bioinformatics tools built for Ion Torrent-generated data

8.3 Overview of advantages and limitations of sequencing approaches

	Pathogen agnostic approaches	**Pathogen specific approaches**
Advantages	Pathogens to be identified do not have to be specified in advance, so sequencing is useful when the likely pathogen identity is unknown, or to identify unusual pathogens present in a sample, or indicate if a novel pathogen is present Can be used to identify multiple microbes of a particular type (e.g. all bacteria or all fungi) by sequencing conserved targets, or microbes of all types using metagenomic methods Can be used to provide context on the microbial community and diversity within a sample, including the presence of co-infections Suitable for direct analysis of samples, without the need for culture. Metagenomic methods sequence all material in the sample, so can be used to identify additional features such as AMR genes, as well as producing data that has the potential to be assembled into partial to whole genomes Targeted methods can be used to enable deeper sequencing of target regions, and pathogen agnostic targets can be combined with pathogen or AMR- specific targets. Sequencing data can be reanalysed if needed, to carry out retrospective investigations into the presence of a particular pathogen	Targeting and/or enriching for specific pathogens increases the amount of genomic material in a sample for sequencing and permits deeper sequencing, increasing data quality and sensitivity of pathogen detection Methods not reliant on culture can be suitable for direct analysis of samples. Methods for WGS result in the most complete genomic data of a single pathogen's sequence, including AMR genes, useful for a range of purposes Different pathogen-specific and AMR gene targets can be combined to identify a range of different pathogens and AMR genes from the same sample. Pathogen specific targets can be used alongside pathogen-agnostic targets May be most appropriate when there is clear need for analysis of just one pathogen of interest, or where there is a high level of confidence in which pathogens should be included on a panel test. Can provide relatively simple, actionable results on whether a particular pathogen or target is present Sequence data from the target regions can be reanalysed to retrospectively identify genetic features of interest

	Pathogen agnostic approaches	Pathogen specific approaches
Limitations	Potential for uncertainty when interpreting which results are clinically relevant to an infection, especially in samples with high levels of commensal microorganisms Can only be used to identify pathogens previously sequenced and where sequence data is present on an accessible database. The quality of the database used will impact upon the quality of pathogen identification The data produced may be in excess to that required and for diagnostic tests may reveal additional results not relevant to the diagnosis Results are easily affected by multiple parameters so standardisation is particularly important Low prevalence of pathogens in samples compared to host DNA can cause challenges, and limit the specificity and sensitivity of pathogen detection It may not be possible to assign genetic features such as AMR genes to specific pathogens detected Assembling a genome from metagenomics can be complex and may result in lower genome coverage than WGS using pathogen specific approaches Pathogen agnostic targeted approaches are not suitable when for pathogens which don't broadly share conserved genome regions e.g. viruses For metagenomics human DNA is also sequenced, so extra requirements over data handling and privacy may apply	Can only be used to identify a defined number of pre-specified pathogens and targets, limiting usability if the identity of the pathogen causing an infection is uncertain. Limited to known pathogens so cannot indicate the presence of novel microorganisms. Genome sequences used in targeted methods may require regular re-evaluation to check they are still suitable as pathogens evolve, and be updated if necessary. Data cannot be reanalysed to retrospectively identify features outside of the target regions, or non-targeted pathogens For non-WGS methods, it is difficult to link AMR genes to pathogen source Limited or no data for understanding microbial community diversity and dynamics within a sample, including co-infections. Culture-based methods are restricted to microorganisms amenable to culture, and require additional facilities and expertise for culture WGS methods can require bioinformatics expertise to assemble genomes and interpret results

8.4 Acknowledgements

We thank the following experts for contributing their time and expertise to project information gathering:

Dr Ahmed Abd el Wahed. Group leader, University of Leipzig, Germany

Dr Silke Arndt. Next generation sequencing service manager, Inqaba Biotechnical Industries, South Africa

Dr Jennifer Bohl. NIH research Fellow at Laboratory of Malaria and Vector Research, US National Institute of Allergy and Infectious Diseases, Phnom Penh, Cambodia

James Brayer. Associate Director, Marketing Development, Oxford Nanopore Technologies, USA

Dr Sónia Gonçalves. Head of Service Delivery, Genomic Surveillance, Wellcome Sanger Institute, UK

Jason Hendry. DPhil student, Oxford University, UK

Alison Howie. Vice President of Commercial, DNAe, UK

Dr Adam Irwin. NHMRC Emerging Leadership Fellow and conjoint Senior Lecturer in paediatric infectious disease, The University of Queensland and Queensland Children's Hospital, Australia

Frédérique Lerêteux. Diagnostic Testing Solutions Manager, Oxford Nanopore Technologies, UK

Prof Iruka Okeke. Professor of Pharmaceutical Microbiology, University of Ibadan, Nigeria

Jenny Overton. Lead consultant, Chartwell consulting, UK

Fionidi Parker. Project Manager, The British Army, UK

Dr Aaron Pomerantz. Segment marketing manager, Oxford Nanopore Technologies, USA

Dr Joshua Quick. UKRI Future Leaders Fellow, University of Birmingham, UK

Dr Senjuti Saha. Director & Scientist, Child Health Research Foundation, Bangladesh

Dr Manu Vanaerschot. Scientist II (PhD), from team of Dr Cristina Tato, Director, Rapid Response Team, Chan Zuckerberg Biohub, USA

Prof David Werner. Professor of Environmental Systems Modelling, University of Newcastle, UK

Dr Amy Wong. Global health consultant, Chan Zuckerberg Initiative.

Dr Christina Yek. Clinical Fellow at Laboratory of Malaria and Vector Research, US National Institute of Allergy and Infectious Diseases, Phnom Penh, Cambodia

We thank the following experts for proving the India MicroLabs case study in section 5.5.1:

Dr Rajesh Pandey. Principal Scientist and PI, CSIR-IGIB and team

9 References

1. Wölfel, R., et al., *Mobile diagnostics in outbreak response, not only for Ebola: a blueprint for a modular and robust field laboratory.* Euro Surveill, 2015. **20**(44): p. pii=30055.
2. Quick, J., et al., *Real-time, portable genome sequencing for Ebola surveillance.* Nature, 2016. **530**(7589): p. 228-232.
3. Jacob, C.G., et al., *Genetic surveillance in the Greater Mekong subregion and South Asia to support malaria control and elimination.* Elife, 2021. **10**: p. e62997.
4. Hansen, S. and A. Abd El Wahed, *Point-Of-Care or Point-Of-Need Diagnostic Tests: Time to Change Outbreak Investigation and Pathogen Detection.* Tropical Medicine and Infectious Disease, 2020. **5**(4): p. 151.
5. Kaplan, C., *Use of the laboratory*, in *Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition.*, H. Walker, W. Hall, and J. Hurst, Editors. 1990, Butterworths: Boston.
6. Wians, F.H., *Clinical Laboratory Tests: Which, Why, and What Do The Results Mean?* Laboratory Medicine, 2009. **40**(2): p. 105-113.
7. Agarwal, R., *Quality-Improvement Measures as Effective Ways of Preventing Laboratory Errors* Laboratory Medicine, 2014. **45**(2): p. e80-e88.
8. Hammerling, J.A., *A Review of Medical Errors in Laboratory Diagnostics and Where We Are Today.* Laboratory Medicine, 2012. **43**(2): p. 41-44.
9. Fonjungo, P.N., et al., *Combatting Global Infectious Diseases: A Network Effect of Specimen Referral Systems.* Clinical Infectious Diseases, 2017. **64**(6): p. 796-803.
10. Kekre, M., et al., *Integrating Scalable Genome Sequencing Into Microbiology Laboratories for Routine Antimicrobial Resistance Surveillance.* Clinical Infectious Diseases, 2021. **73**(S4): p. S258-S266.
11. Chen, H., et al., *Point of care testing for infectious diseases.* Clinica Chimica Acta, 2019. **493**: p. 138-147.
12. Kettler, H., K. White, and S.J. Hawkes. *Mapping the landscape of diagnostics for sexually transmitted infections : key findings and recommendations*. 2004; Available from: https://apps.who.int/iris/handle/10665/68990.
13. Sachdeva, S., R.W. Davis, and A.K. Saha, *Microfluidic Point-of-Care Testing: Commercial Landscape and Future Directions.* Frontiers in Bioengineering and Biotechnology, 2021. **8**: p. 602659.
14. World Health Organization. *Nucleic acid amplification-based diagnostics*. 2022; Available from: https://www.who.int/teams/global-malaria-programme/case-management/diagnosis/nucleic-acid-amplification-based-diagnostics.
15. Holland, C.A. and F.L. Kiechle, *Point-of-care molecular diagnostic systems — past, present and future.* Current Opinion in Microbiology, 2005. **8**(5): p. 504-509.
16. Zarei, M., *Advances in point-of-care technologies for molecular diagnostics.* Biosensors and Bioelectronics, 2017. **98**: p. 494-506.
17. Tokel, O., F. Inci, and U. Demirci, *Advances in plasmonic technologies for point of care applications.* Chemical Reviews, 2014. **114**(11): p. 5728-5752.
18. Hsieh, H.V., J.L. Dantzler, and B.H. Weigl, *Analytical Tools to Improve Optimization Procedures for Lateral Flow Assays.* Diagnostics, 2017. **7**(2): p. 29.
19. Escadafal, C., et al., *New Biomarkers and Diagnostic Tools for the Management of Fever in Low- and Middle-Income Countries: An Overview of the Challenges.* Diagnostics, 2017. **7**(3): p. 44.
20. Neher, R.A. and T. Bedford, *Real-Time Analysis and Visualization of Pathogen Sequence Data.* Journal of Clinical Microbiology, 2018. **56**(11): p. e00480-18.
21. MacLean, E., et al., *Advances in Molecular Diagnosis of Tuberculosis.* Journal of Clinical Microbiology, 2020. **58**(10): p. e01582-19.
22. Schwarze, K., et al., *The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom.* Genetics in Medicine, 2020. **22**(1): p. 85-94.
23. Smith, D.I. *Front Line Genomics: Sequencing Buyers Guide, 3rd Edition*. 2021; Available from: https://lp.frontlinegenomics.com/sequencing-buyers-guide-3rd-edition.
24. Krehenwinkel, H., A. Pomerantz, and S. Prost, *Genetic Biomonitoring and Biodiversity Assessment Using Portable Sequencing Technologies: Current Uses and Future Directions.* Genes, 2019. **10**(11): p. 858.

25. World Health Organization. *GLASS whole-genome sequencing for surveillance of antimicrobial resistance*. 2020; Available from: https://www.who.int/publications/i/item/9789240011007.
26. Oxford Nanopore Technologies. *MinION*. 2022; Available from: https://nanoporetech.com/products/minion.
27. Oxford Nanopore Technologies. *Flongle*. 2022; Available from: https://nanoporetech.com/products/flongle.
28. Vereecke, N., et al., *High quality genome assemblies of Mycoplasma bovis using a taxon-specific Bonito basecaller for MinION and Flongle long-read nanopore sequencing.* BMC Bioinformatics, 2020. **21**(1): p. 517.
29. Marcolungo, L., et al., *Real-Time On-Site Diagnosis of Quarantine Pathogens in Plant Tissues by Nanopore-Based Sequencing.* Pathogens, 2022. **11**(2): p. 199.
30. Castro-Wallace, S.L., et al., *Nanopore DNA Sequencing and Genome Assembly on the International Space Station.* Scientific Reports, 2017. **7**(1): p. 18022.
31. Johnson, S.S., et al., *Real-Time DNA Sequencing in the Antarctic Dry Valleys Using the Oxford Nanopore Sequencer.* Journal of Biomolecular Techniques, 2017. **28**(1): p. 2-7.
32. Pomerantz, A., et al., *Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building.* GigaScience, 2018. **7**(4): p. giy033.
33. Latorre-Pérez, A., et al., *A Round Trip to the Desert: In situ Nanopore Sequencing Informs Targeted Bioprospecting.* Frontiers in Microbiology, 2021. **12**: p. 768240.
34. Reintjes, G., et al., *On-Site Analysis of Bacterial Communities of the Ultraoligotrophic South Pacific Gyre.* Applied and Environmental Microbiology, 2019. **85**(14): p. e00184-19.
35. Gowers, G.-O.F., et al., *Entirely Off-Grid and Solar-Powered DNA Sequencing of Microbial Communities during an Ice Cap Traverse Expedition.* Genes, 2019. **10**(11): p. 902.
36. Hansen, S., et al., *Serotyping of foot-and-mouth disease virus using Oxford nanopore sequencing.* Journal of Virological Methods, 2019. **263**: p. 50-53.
37. Rambo-Martin, B.L., et al., *Influenza A Virus Field Surveillance at a Swine-Human Interface.* mSphere, 2020. **5**(1): p. e00822-19.
38. Zhang, Q., et al., *Application of metagenomic next-generation sequencing (mNGS) combined with rapid on-site cytological evaluation (ROSCE) for the diagnosis of Chlamydia psittaci pneumonia.* International Journal of Clinical and Experimental Pathology, 2021. **14**(4): p. 389-398.
39. Votintseva, A.A., et al., *Same-Day Diagnostic and Surveillance Data for Tuberculosis via Whole-Genome Sequencing of Direct Respiratory Samples.* Journal of Clinical Microbiology, 2017. **55**(5): p. 1285-1298.
40. Hirabayashi, A., et al., *On-Site Genomic Epidemiological Analysis of Antimicrobial-Resistant Bacteria in Cambodia With Portable Laboratory Equipment.* Frontiers in Microbiology, 2021. **12**: p. 675463.
41. Fu, S., A. Wang, and K.F. Au, *A comparative evaluation of hybrid error correction methods for error-prone long reads.* Genome Biology, 2019. **20**(1): p. 26.
42. Amarasinghe, S.L., et al., *Opportunities and challenges in long-read sequencing data analysis.* Genome Biology, 2020. **21**(1): p. 30.
43. Acharya, K., et al., *Metagenomic water quality monitoring with a portable laboratory.* Water Research, 2020. **184**: p. 116112.
44. Stefan, C.P., et al., *Comparison of Illumina and Oxford Nanopore Sequencing Technologies for Pathogen Detection from Clinical Matrices Using Molecular Inversion Probes.* The Journal of Molecular Diagnostics, 2022. **4**: p. 395-405.
45. Oxford Nanopore Technologies. *VolTRAX*. 2022; Available from: https://nanoporetech.com/products/voltrax.
46. Oxford Nanopore Technologies. *SmidgION*. 2022; Available from: https://nanoporetech.com/products/smidgion.
47. Oxford Nanopore Technologies. *MinION Mk1D*. 2022; Available from: https://nanoporetech.com/products/minion-mk1d.
48. Oxford Nanopore Technologies. *PromethION P2*. 2022; Available from: https://nanoporetech.com/products/p2.
49. Illumina. *Illumina delivers 2021 results ahead of expectations, strong 2022 guidance and deep pipeline of long-term growth opportunities at 2022 J.P. Morgan Healthcare Conference*. 2022; Available from: https://emea.illumina.com/company/news-center/press-releases/press-release-details.html?newsid=e3ad6d4e-fcbe-4a0d-ac1d-5a3996d1bfd8.
50. Illumina. *NextSeq 550Dx*. 2022; Available from: https://emea.illumina.com/systems/sequencing-platforms/nextseq-dx.html.

51. Mbala-Kingebeni, P., et al., *2018 Ebola virus disease outbreak in Équateur Province, Democratic Republic of the Congo: a retrospective genomic characterisation.* The Lancet Infectious Diseases, 2019. **19**(6): p. 641-647.
52. Illumina. *iSeq 100 system* 2022; Available from: https://emea.illumina.com/systems/sequencing-platforms/iseq.html.
53. Colman, R.E., et al., *Whole-genome and targeted sequencing of drug-resistant Mycobacterium tuberculosis on the iSeq100 and MiSeq: A performance, ease-of-use, and cost evaluation.* PLoS Medicine, 2019. **16**(4): p. e1002794.
54. Nakao, R., et al., *Illumina iSeq 100 and MiSeq exhibit similar performance in freshwater fish environmental DNA metabarcoding.* Scientific Reports, 2021. **11**(1): p. 15763.
55. Bohl, J.A., et al., *Discovering disease-causing pathogens in resource-scarce Southeast Asia using a global metagenomic pathogen monitoring system.* Proceedings of the National Academy of Sciences, 2022. **119**(11): p. e2115285119.
56. Kinganda-Lusamaki, E., et al., *Integration of genomic sequencing into the response to the Ebola virus outbreak in Nord Kivu, Democratic Republic of the Congo.* Nature Medicine, 2021. **27**(4): p. 710-716.
57. Thermo Fisher Scientific. *Ion Torrent Genexus System*. 2022; Available from: https://www.thermofisher.com/uk/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-workflow/ion-torrent-next-generation-sequencing-run-sequence/ion-torrent-genexus-system.html.
58. Ion Torrent. *Ion Chef System specification sheet*. 2017; Available from: http://assets.thermofisher.com/TFS-Assets/LSG/Application-Notes/Ion_Chef_Specification_Sheet.PDF.
59. Arias, A., et al., *Rapid outbreak sequencing of Ebola virus in Sierra Leone identifies transmission chains linked to sporadic cases.* Virus Evolution, 2016. **2**(1): p. vew016.
60. Hansen, S., et al., *Combination random isothermal amplification and nanopore sequencing for rapid identification of the causative agent of an outbreak.* Journal of Clinical Virology, 2018. **106**: p. 23-27.
61. Mbala-Kingebeni, P., et al., *Medical countermeasures during the 2018 Ebola virus disease outbreak in the North Kivu and Ituri Provinces of the Democratic Republic of the Congo: a rapid genomic assessment.* The Lancet Infectious Diseases, 2019. **19**(6): p. 648-657.
62. Hourdel, V., et al., *Rapid Genomic Characterization of SARS-CoV-2 by Direct Amplicon-Based Sequencing Through Comparison of MinION and Illumina iSeq100TM System.* Frontiers in Microbiology, 2020. **11**: p. 571328.
63. Razook, Z., et al., *Real time, field-deployable whole genome sequencing of malaria parasites using nanopore technology.* bioRxiv, 2020: p. 2020.12.17.423341.
64. Gwinn, M., D. MacCannell, and G.L. Armstrong, *Next-Generation Sequencing of Infectious Pathogens.* JAMA, 2019. **321**(9): p. 893-894.
65. Tshiabuila, D., et al., *Comparison of SARS-CoV-2 sequencing using the ONT GridION and the Illumina MiSeq.* BMC Genomics, 2022. **23**(1): p. 319.
66. Heikema, A.P., et al., *Comparison of Illumina versus Nanopore 16S rRNA Gene Sequencing of the Human Nasal Microbiota.* Genes, 2020. **11**(9).
67. GenomeWeb. *Roche's 454 Sues Thermo Fisher's Ion Torrent for Patent Infringement*. 2015; Available from: https://www.genomeweb.com/business-news/roches-454-sues-thermo-fishers-ion-torrent-patent-infringement#.YjNCL3rP0dU.
68. GenomeWeb. *Illumina Wins $8M Jury Verdict in BGI Patent Infringement Suit, Loses Patent*. 2021; Available from: https://www.genomeweb.com/sequencing/illumina-wins-8m-jury-verdict-bgi-patent-infringement-suit-loses-patent#.YjNEDXrP0dV.
69. Front Line Genomics. *The Sequencing Buyer's Guide, 4th Edition* 2022; Available from: https://frontlinegenomics.com/the-sequencing-buyers-guide-4th-edition/.
70. Kafetzopoulou, L.E., et al., *Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak.* Science, 2019. **363**(6422): p. 74-77.
71. Simner, P.J., S. Miller, and K.C. Carroll, *Understanding the Promises and Hurdles of Metagenomic Next-Generation Sequencing as a Diagnostic Tool for Infectious Diseases.* Clin Infect Dis, 2018. **66**(5): p. 778-788.
72. Ursenbach, A., et al., *Impact of 16S rDNA sequencing on clinical treatment decisions: a single center retrospective study.* BMC Infectious Diseases, 2021. **21**(1): p. 190.
73. Great Ormond Street Hospital for Children. *16S rDNA PCR*. 2022; Available from: http://www.labs.gosh.nhs.uk/laboratory-services/microbiology-virology-and-infection-control/tests/16s-rdna-pcr.

74. Aggarwal, D., et al., *Clinical utility and cost-effectiveness of bacterial 16S rRNA and targeted PCR based diagnostic testing in a UK microbiology laboratory network.* Scientific Reports, 2020. **10**(1): p. 7965.
75. DNAe. *Sepsis*. 2022; Available from: https://www.dnae.com/sepsis.
76. Illumina. *Respiratory Pathogen ID/AMR Enrichment Panel Kit (Powered by IDbyDNA Explify)*. 2022; Available from: https://sapac.illumina.com/products/by-type/sequencing-kits/library-prep-kits/respiratory-pathogen-id-panel.html.
77. FIND. *Seq&Treat project*. 2022; Available from: https://www.finddx.org/at-risk-populations/seq-treat/.
78. World Health Organization. *Technical manual for drug susceptibility testing of medicines used in the treatment of tuberculosis.* 2018; Available from: https://apps.who.int/iris/handle/10665/275469.
79. Zhang, H. *FIND Project Aims to Accelerate Targeted NGS Solutions for Global Drug-Resistant TB Testing*. 2022; Available from: https://www.genomeweb.com/clinical-sequencing/find-project-aims-accelerate-targeted-ngs-solutions-global-drug-resistant-tb#.YmifnNrMJPY.
80. Artic Network. *Protocols* 2022; Available from: https://artic.network/2-protocols.html.
81. Quick, J., et al., *Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples.* Nature Protocols, 2017. **12**(6): p. 1261-1276.
82. Beukers, A.G., F. Jenkins, and S.J. van Hal, *Centralised or Localised Pathogen Whole Genome Sequencing: Lessons Learnt From Implementation in a Clinical Diagnostic Laboratory.* Frontiers in Cellular and Infection Microbiology, 2021. **11**: p. 636290.
83. Duan, H., et al., *The diagnostic value of metagenomic next-generation sequencing in infectious diseases.* BMC Infectious Diseases, 2021. **21**(1): p. 62.
84. World Health Organization. *Laboratory biosafety manual, 4th edition*. 2020; Available from: https://www.who.int/publications/i/item/9789240011311.
85. UN3373. *Regulations for UN3373*. 2022; Available from: https://www.un3373.com/enveloppeninfo/regulations.
86. Robin, J.D., et al., *Comparison of DNA Quantification Methods for Next Generation Sequencing.* Scientific Reports, 2016. **6**(1): p. 24067.
87. Ramos-Mandujano, G., et al., *A Robust, Safe, and Scalable Magnetic Nanoparticle Workflow for RNA Extraction of Pathogens from Clinical and Wastewater Samples.* Global Challenges, 2021. **5**(4): p. 2000068.
88. Hansen, S., et al., *Zika Virus Amplification Using Strand Displacement Isothermal Method and Sequencing Using Nanopore Technology*, in *Zika Virus: Methods and Protocols*, G. Kobinger and T. Racine, Editors. 2020, Springer US: New York, NY. p. 123-136.
89. Li, J., J. Macdonald, and F. von Stetten, *Review: a comprehensive summary of a decade development of the recombinase polymerase amplification.* Analyst, 2019. **144**(1): p. 31-67.
90. MiniPCR. *miniPCR® mini8 thermal cycler*. 2022; Available from: https://www.minipcr.com/products/minipcr/.
91. González-González, E., et al., *Validation of use of the miniPCR thermocycler for Ebola and Zika virus detection.* PLoS ONE, 2019. **14**(5): p. e0215642.
92. Thermo Fisher Scientific. *The Long and Short of Isothermal Amplification*. 2022 [cited 2022; Available from: https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/spotlight-articles/isothermal-amplification.html.
93. Oliveira, B.B., B. Veigas, and P.V. Baptista, *Isothermal Amplification of Nucleic Acids: The Race for the Next "Gold Standard".* Frontiers in Sensors, 2021. **2**: p. 10.3389/fsens.2021.752600.
94. Ptasinska, A., et al., *Diagnostic accuracy of loop-mediated isothermal amplification coupled to nanopore sequencing (LamPORE) for the detection of SARS-CoV-2 infection at scale in symptomatic and asymptomatic populations.* Clinical Microbiology and Infection, 2021. **27**(9): p. 1348.e1-1348.e7.
95. Martin, S., et al., *Nanopore adaptive sampling: a tool for enrichment of low abundance species in metagenomic samples.* Genome Biology, 2022. **23**(1): p. 11.
96. Jayatilleke, K., *Challenges in Implementing Surveillance Tools of High-Income Countries (HICs) in Low Middle Income Countries (LMICs).* Curr Treat Options Infect Dis, 2020: p. 1-11.
97. Lau, K.A., et al., *Proficiency testing for bacterial whole genome sequencing in assuring the quality of microbiology diagnostics in clinical and public health laboratories.* Pathology, 2021. **53**(7): p. 902-911.
98. Trivedi, U.H., et al., *Quality control of next-generation sequencing data without a reference.* Frontiers in Genetics, 2014. **5**: p. 111.

99. Kozyreva Varvara, K., et al., *Validation and Implementation of Clinical Laboratory Improvements Act-Compliant Whole-Genome Sequencing in the Public Health Microbiology Laboratory.* Journal of Clinical Microbiology, 2017. **55**(8): p. 2502-2520.
100. Walton, R.M., *Validation of laboratory tests and methods.* Seminars in Avian and Exotic Pet Medicine, 2001. **10**(2): p. 59-65.
101. Public Health England. *Proficiency testing: the process and benefits*. 2015; Available from: https://www.gov.uk/guidance/external-quality-assessment-eqa-and-proficiency-testing-pt-the-process-and-benefits.
102. Racine, T. and G.P. Kobinger, *Challenges and perspectives on the use of mobile laboratories during outbreaks and their use for vaccine evaluation.* Human Vaccines & Immunotherapeutics, 2019. **15**(10): p. 2264-2268.
103. Goodfellow, I., C. Reusken, and M. Koopmans, *Laboratory support during and after the Ebola virus endgame: towards a sustained laboratory infrastructure.* Eurosurveillance, 2015. **20**(12): p. 21074.
104. World Health Organization Regional Office for Europe. *Guidance For Rapid Response Mobile Laboratory (RRML) Classification*. 2021; Available from: https://apps.who.int/iris/bitstream/handle/10665/339845/9789289054928-eng.pdf.
105. Mackenzie, J.S., et al., *The global outbreak alert and response network.* Glob Public Health, 2014. **9**(9): p. 1023-39.
106. Weidmann, M., et al., *Development of Mobile Laboratory for Viral Hemorrhagic Fever Detection in Africa.* The Journal of Infectious Diseases, 2018. **218**(10): p. 1622-1630.
107. Affara, M., et al., *The East African Community (EAC) mobile laboratory networks in Kenya, Burundi, Tanzania, Rwanda, Uganda, and South Sudan—from project implementation to outbreak response against Dengue, Ebola, COVID-19, and epidemic-prone diseases.* BMC Medicine, 2021. **19**(1): p. 160.
108. Subissi, L., et al., *A large epidemic of a necrotic skin infection in the Democratic Republic of São Tomé and Principe: an epidemiological study.* International Journal of Infectious Diseases, 2021. **110**: p. S69-S76.
109. Brunker, K., et al., *Rapid in-country sequencing of whole virus genomes to inform rabies elimination programmes.* Wellcome Open Research, 2020. **5**: p. 3.
110. capacity4dev. *European Mobile Laboratory Project - Stephan Günther*. 2014; Available from: https://www.youtube.com/watch?v=xv79ak1xv80.
111. Capacity4Dev. *EU Mobile Labs: Part of the Solution to Tackle the Ebola Outbreak*. 2014; Available from: https://europa.eu/capacity4dev/articles/eu-mobile-labs-part-solution-tackle-ebola-outbreak.
112. East African Community. *Phase II Container Laboratories*. 2022; Available from: https://www.eac.int/project-milestones/mobile-laboratories/238-sector/health/mobile-labs.
113. Wellcome Trust. *Lab in a suitcase: mobile genetic sequencing for outbreak response*. 2018; Available from: https://www.youtube.com/watch?v=sjwZMMFZhRY.
114. Quick, J. *How to build a mobile lab for genomic surveillance of viral outbreaks (Video)*. 2021; Available from: https://www.myeventflo.com/event-lecture.asp?m=0&evID=2374&lectID=24573&list=2.
115. Walker, K.T., et al., *CONTAIN: An open-source shipping container laboratory optimised for automated COVID-19 diagnostics.* bioRxiv, 2020: p. 2020.05.20.106625.
116. Bilooei, S.F., et al., *Rapid genome surveillance of SARS-CoV-2 and study of risk factors using shipping container laboratories and portable DNA sequencing technology.* medRxiv, 2022: p. 10.1101/2022.02.25.22271277.
117. Baraniuk, C., *A covid-19 laboratory for Jersey—in a shipping container.* BMJ, 2020. **370**: p. m3336.
118. itvNews. *Lab closure to mean longer waits for pre-departure Covid-19 tests in Jersey*. 2021; Available from: https://www.itv.com/news/channel/2021-11-10/lab-closure-to-mean-longer-waits-for-pre-departure-covid-19-tests-in-jersey.
119. HL7 UK. *HL7 delivers healthcare interoperability standards*. 2022; Available from: https://www.hl7.org.uk/.
120. National Health Service (UK). *National Pathology FHIR Messaging Specifications*. 2022; Available from: https://developer.nhs.uk/apis/itk3nationalpathology-1-1-0/.
121. Ramanathan, A. *The Mobile Laboratory: How a Scientist Tests for Infectious Diseases out of a Suitcase*. 2017; Available from: https://www.selectscience.net/editorial-articles/the-mobile-laboratory-how-a-scientist-tests-for-infectious-diseases-out-of-a-suitcase/?artID=43898.
122. Oxford Nanopore Technologies. *Webinar: Nanopore sequencing: the missing puzzle piece in molecular identification. Ahmed Abd El Wahed*. 2022; Available from:

https://nanoporetech.com/resource-centre/nanopore-sequencing-missing-puzzle-piece-molecular-identification.
123. Illumina. *Bringing Sophisticated Lab Experiences to High School Classrooms*. 2021; Available from: https://www.illumina.com/company/news-center/feature-articles/lab-experiences-in-high-school-classrooms.html.
124. Royal Geographical Society. *Mobile Malaria project update*. 2019; Available from: https://www.rgs.org/geography/news/mobile-malaria-project-update/.
125. Mobile Malaria Project. *Can you get a lab into the boot of a Land Rover?* 2019; Available from: https://mobilemalaria.com/about/lab-rover/.
126. Zibra project. *Zika in Brazil real time analysis*. 2016; Available from: http://www.zibraproject.org/mobile/.
127. IMechE Yorkshire Region. *Engineering Out Of A Pandemic - How Engineering Supported A Clinical Problem*. 2022; Available from: https://www.youtube.com/watch?v=Lu8T7nwyBIE.
128. ARTIC Network. *Ebola Virus*. 2022; Available from: https://artic.network/ebov.
129. Watsa, M., et al., *Portable sequencing as a teaching tool in conservation and biodiversity research.* PLoS Biology, 2020. **18**(4): p. e3000667.
130. Singer, M., et al., *The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3).* JAMA, 2016. **315**(8): p. 801-10.
131. Schultz, M.J., et al., *Current Challenges in the Management of Sepsis in ICUs in Resource-Poor Settings and Suggestions for the Future*, in *Sepsis Management in Resource-limited Settings*, A.M. Dondorp, M.W. Dünser, and M.J. Schultz, Editors. 2019, Springer Copyright 2019, The Author(s). Cham (CH). p. 1-24.
132. Rudd, K.E., et al., *Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study.* The Lancet, 2020. **395**(10219): p. 200-211.
133. *Sepsis Management in Resource-limited Settings*, ed. A.M. Dondorp, Dünser, M.W., Schultz, M.J. 2019: Springer, Cham.
134. World Health Organization. *Sepsis*. 2020; Available from: https://www.who.int/news-room/fact-sheets/detail/sepsis.
135. Thwaites, C.L., et al., *Infection Management in Patients with Sepsis and Septic Shock in Resource-Limited Settings*, in *Sepsis Management in Resource-limited Settings*, A.M. Dondorp, M.W. Dünser, and M.J. Schultz, Editors. 2019, Springer Copyright 2019, The Author(s). Cham (CH). p. 163-84.
136. Moraes, R.B., et al., *De-escalation, adequacy of antibiotic therapy and culture positivity in septic patients: an observational study.* Revista Brasileira de Terapia Intensiva, 2016. **28**(3): p. 315-322.
137. Martínez, M.L., et al., *An approach to antibiotic treatment in patients with sepsis.* Journal of Thoracic Disease, 2020. **12**(3): p. 1007-1021.
138. Rhee, C., et al., *Prevalence of Antibiotic-Resistant Pathogens in Culture-Proven Sepsis and Outcomes Associated With Inadequate and Broad-Spectrum Empiric Antibiotic Use.* JAMA Network Open, 2020. **3**(4): p. e202899-e202899.
139. Evans, L., et al., *Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021.* Intensive Care Medicine, 2021. **47**(11): p. 1181-1247.
140. UK Health Security Agency. *UK Standards for Microbiology Investigations Sepsis, and other systemic and disseminated infections*. 2021; Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1037444/S_12_dzj+.pdf.
141. Flannery, D.D., et al., *Delivery Characteristics and the Risk of Early-Onset Neonatal Sepsis.* Pediatrics, 2022. **149**(2): p. e2021052900.
142. BioFire Diagnostics. *The BioFire® FilmArray® Blood Culture Identification Panels*. 2022; Available from: https://www.biofiredx.com/products/the-filmarray-panels/filmarraybcid/.
143. Chun, K., et al., *Sepsis Pathogen Identification.* J Lab Autom, 2015. **20**(5): p. 539-61.
144. T2 Biosystems. *T2Bacteria Panel*. 2021; Available from: https://www.t2biosystems.com/products-technology-ous/t2bacteria-panel-ous/.
145. Global AMR Hub. *Estimating global patient needs and market potential for priority health technologies addressing antimicrobial resistance*. 2021; Available from: https://globalamrhub.org/wp-content/uploads/2021/08/EAG-Report_FINAL_20082021.pdf.
146. MeMed. *MeMed BV*. 2022; Available from: https://www.me-med.com/.
147. Thomson, K.M., et al., *Effects of antibiotic resistance, drug target attainment, bacterial pathogenicity and virulence, and antibiotic access and affordability on outcomes in neonatal sepsis: an international microbiology and drug evaluation prospective substudy (BARNARDS).* Lancet Infect Dis, 2021. **21**(12): p. 1677-1688.

148. Rozo, M., et al., *An Observational Study of Sepsis in Takeo Province Cambodia: An in-depth examination of pathogens causing severe infections.* PLoS Negl Trop Dis, 2020. **14**(8): p. e0008381.
149. Sinha, M., et al., *Emerging Technologies for Molecular Diagnosis of Sepsis.* Clinical Microbiology Reviews, 2018. **31**(2): p. e00089-17.
150. Mellhammar, L., et al., *Bacteremic sepsis leads to higher mortality when adjusting for confounders with propensity score matching.* Scientific Reports, 2021. **11**(1): p. 6972.
151. Ombelet, S., et al., *Best Practices of Blood Cultures in Low- and Middle-Income Countries.* Frontiers in Medicine, 2019. **6**: p. 131.
152. Clinical trial. *Clinical Impact of Rapid Identification of Positive Blood Cultures vs. Internal Laboratory Standard*. 2022; Available from: https://clinicaltrials.gov/ct2/show/NCT04156633.
153. World Health Organization. *Global report on the epidemiology and burden of sepsis* 2020; Available from: https://www.who.int/publications/i/item/9789240010789.
154. Li, J.-Y., et al., *Nanopore-targeted sequencing for simultaneous diagnosis of suspected sepsis and early targeted therapy.* Annals of Translational Medicine, 2021. **9**(23): p. 1749.
155. Grumaz, S., et al., *Enhanced Performance of Next-Generation Sequencing Diagnostics Compared With Standard of Care Microbiological Diagnostics in Patients Suffering From Septic Shock.* Crit Care Med, 2019. **47**(5): p. e394-e402.
156. Clinical trial. *Next-Generation Sequencing Diagnostics of Bacteremia in Sepsis (NextGeneSiS)*. 2022; Available from: https://clinicaltrials.gov/ct2/show/NCT03356249.
157. Schmoch, T., et al., *Next-generation sequencing diagnostics of bacteremia in pediatric sepsis.* Medicine, 2021. **100**(25): p. e26403.
158. Brenner, T., et al., *Optimization of sepsis therapy based on patient-specific digital precision diagnostics using next generation sequencing (DigiSep-Trial)—study protocol for a randomized, controlled, interventional, open-label, multicenter trial.* Trials, 2021. **22**(1): p. 714.
159. Irwin, A.D., et al., *Optimising Treatment Outcomes for Children and Adults Through Rapid Genome Sequencing of Sepsis Pathogens. A Study Protocol for a Prospective, Multi-Centre Trial (DIRECT).* Frontiers in Cellular and Infection Microbiology, 2021. **11**: p. 667680.
160. Karius Diagnostics. *The Karius Test*. 2022; Available from: https://kariusdx.com/the-karius-test/.
161. Babady, N.E., *Clinical Metagenomics for Bloodstream Infections: Is the Juice Worth the Squeeze?* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2021. **72**(2): p. 246-248.
162. PathoQuest. *iDTECT Dx test. A new approach to pathogen identification*. 2022; Available from: https://previous.pathoquest.com/?page_id=69.
163. Moragues-Solanas, L., R. Scotti, and J. O'Grady, *Rapid metagenomics for diagnosis of bloodstream and respiratory tract nosocomial infections: current status and future prospects.* Expert Review of Molecular Diagnostics, 2021. **21**(4): p. 371-380.
164. Parize, P., 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**(8): p. 574.e1-574.e6.
165. Molzym. *Molecular Pathogen Diagnosis: SepsiTest™-UMD CE IVD*. 2022; Available from: https://www.molzym.com/pathogen-diagnosis.
166. Fraunhofer Institute for Cell Therapy and Immunology. *AutoSepT – Automated sepsis diagnostic for fast testing in clinical laboratory*. 2020; Available from: https://www.izi-bb.fraunhofer.de/en/media/press/2020/autosept.html.
167. Pathogenomix. *Patho-Seq*. 2021; Available from: https://www.pathogenomix.com/pathogenomix_pathoseq.aspx.
168. *Day Zero Diagnostics*. 2021; Available from: https://www.dayzerodiagnostics.com/.
169. Noviyanti, R., et al., *Implementing parasite genotyping into national surveillance frameworks: feedback from control programmes and researchers in the Asia-Pacific region.* Malaria Journal, 2020. **19**(1): p. 271.
170. Ferkol, T. and D. Schraufnagel, *The global burden of respiratory disease.* Annals of the American Thoracic Society, 2014. **11**(3): p. 404-406.
171. World Health Organization. *The top 10 causes of death*. 2020; Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
172. Troeger, C., et al., *Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015.* The Lancet Infectious Diseases, 2017. **17**(11): p. 1133-1161.
173. Savitha, M., et al., *Modifiable risk factors for acute lower respiratory tract infections.* The Indian Journal of Pediatrics, 2007. **74**(5): p. 477-482.

174. Rali, P., et al., *Opportunistic pulmonary infections in immunocompromised hosts.* Critical Care Nursing Quarterly, 2016. **39**(2): p. 161-175.
175. World Health Organization. *2019 antibacterial agents in clinical development: an analysis of the antibacterial clinical development pipeline*. 2019; Available from: https://www.who.int/publications/i/item/9789240000193.
176. Rytter, H., et al., *Which Current and Novel Diagnostic Avenues for Bacterial Respiratory Diseases?* Frontiers in Microbiology, 2020. **11**: p. 616971.
177. James, S.L., et al., *Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017.* The Lancet, 2018. **392**(10159): p. 1789-1858.
178. Roth, G.A., et al., *Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017.* The Lancet, 2018. **392**(10159): p. 1736-1788.
179. Fendrick, A.M., et al., *The economic burden of non–influenza-related viral respiratory tract infection in the United States.* Archives of Internal Medicine, 2003. **163**(4): p. 487-494.
180. Brown, G.D., et al., *Hidden Killers: Human Fungal Infections.* Science Translational Medicine, 2012. **4**(165): p. 165rv13.
181. McTaggart, L.R., et al., *Mycobiome Sequencing and Analysis Applied to Fungal Community Profiling of the Lower Respiratory Tract During Fungal Pathogenesis.* Frontiers in Microbiology, 2019. **10**: p. 512.
182. de Dios Caballero, J., et al., *The Human Mycobiome in Chronic Respiratory Diseases: Current Situation and Future Perspectives.* Microorganisms, 2022. **10**(4): p. 810.
183. Couturier, M.R., E.H. Graf, and A.T. Griffin, *Urine antigen tests for the diagnosis of respiratory infections: legionellosis, histoplasmosis, pneumococcal pneumonia.* Clin Lab Med, 2014. **34**(2): p. 219-36.
184. Garcin, F., et al., *Non-adherence to guidelines: an avoidable cause of failure of empirical antimicrobial therapy in the presence of difficult-to-treat bacteria.* Intensive Care Med, 2010. **36**(1): p. 75-82.
185. Lim, W.S., et al., *BTS guidelines for the management of community acquired pneumonia in adults: update 2009.* Thorax, 2009. **64**(S3): p. iii1.
186. Jain, S., et al., *Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults.* New England Journal of Medicine, 2015. **373**(5): p. 415-427.
187. Diao, Z., et al., *Metagenomics next-generation sequencing tests take the stage in the diagnosis of lower respiratory tract infections.* Journal of Advanced Research, 2021(38): p. 201-212.
188. Tenover, F.C., *Developing molecular amplification methods for rapid diagnosis of respiratory tract infections caused by bacterial pathogens.* Clin Infect Dis, 2011. **52**(S4): p. S338-45.
189. Charalampous, T., et al., *Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection.* Nature Biotechnology, 2019. **37**(7): p. 783-792.
190. Bosshard, P.P., *Incubation of fungal cultures: how long is long enough?* Mycoses, 2011. **54**(5): p. e539-e545.
191. Balajee, S.A., et al., *Sequence-Based Identification of Aspergillus, Fusarium, and Mucorales Species in the Clinical Mycology Laboratory: Where Are We and Where Should We Go from Here?* Journal of Clinical Microbiology, 2009. **47**(4): p. 877-884.
192. Schoch, C.L., et al., *Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi.* Proceedings of the National Academy of Sciences, 2012. **109**(16): p. 6241-6246.
193. von Laer, A., et al., *Implementation of Automated Blood Culture With Quality Assurance in a Resource-Limited Setting.* Frontiers in Medicine, 2021. **8**: p. 627513-627513.
194. Cookson, W.O.C.M., M.J. Cox, and M.F. Moffatt, *New opportunities for managing acute and chronic lung infections.* Nature Reviews Microbiology, 2018. **16**(2): p. 111-120.
195. Nelson, P.P., et al., *Current and Future Point-of-Care Tests for Emerging and New Respiratory Viruses and Future Perspectives.* Frontiers in Cellular and Infection Microbiology, 2020. **10**: p. 181.
196. Das, S., S. Dunbar, and Y.-W. Tang, *Laboratory Diagnosis of Respiratory Tract Infections in Children – the State of the Art.* Frontiers in Microbiology, 2018. **9**: p. 2478.
197. Větrovský, T., et al., *GlobalFungi, a global database of fungal occurrences from high-throughput-sequencing metabarcoding studies.* Scientific Data, 2020. **7**(1): p. 228.
198. Janda, J.M. and S.L. Abbott, *16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls.* J Clin Microbiol, 2007. **45**(9): p. 2761-4.

199. Woo, P.C.Y., et al., *Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories.* Clinical Microbiology and Infection, 2008. **14**(10): p. 908-934.
200. Clarridge, J.E., 3rd, *Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases.* Clinical Microbiology Reviews, 2004. **17**(4): p. 840-862.
201. López-Labrador, F.X., et al., *Recommendations for the introduction of metagenomic high-throughput sequencing in clinical virology, part I: Wet lab procedure.* Journal of Clinical Virology, 2021. **134**: p. 104691.
202. Xie, F., et al., *Clinical metagenomics assessments improve diagnosis and outcomes in community-acquired pneumonia.* BMC Infectious Diseases, 2021. **21**(1): p. 352.
203. Zhu, N., et al., *A Novel Coronavirus from Patients with Pneumonia in China, 2019.* New England Journal of Medicine, 2020. **382**(8): p. 727-733.
204. Liu, D., et al., *Multicenter assessment of shotgun metagenomics for pathogen detection.* eBioMedicine, 2021. **74**: p. 103649.
205. Zheng, Y., et al., *The Diagnostic Value of Metagenomic Next–Generation Sequencing in Lower Respiratory Tract Infection.* Frontiers in Cellular and Infection Microbiology, 2021. **11**: p. 694756.
206. Langelier, C., et al., *Integrating host response and unbiased microbe detection for lower respiratory tract infection diagnosis in critically ill adults.* Proc Natl Acad Sci U S A, 2018. **115**(52): p. E12353-E12362.
207. Wang, J., Y. Han, and J. Feng, *Metagenomic next-generation sequencing for mixed pulmonary infection diagnosis.* BMC Pulmonary Medicine, 2019. **19**(1): p. 252.
208. van Rijn, A.L., et al., *The respiratory virome and exacerbations in patients with chronic obstructive pulmonary disease.* PLoS ONE, 2019. **14**(10): p. e0223952.
209. Li, H., et al., *Detection of Pulmonary Infectious Pathogens From Lung Biopsy Tissues by Metagenomic Next-Generation Sequencing.* Front Cell Infect Microbiol, 2018. **8**: p. 205.
210. Huang, J., et al., *Metagenomic next-generation sequencing versus traditional pathogen detection in the diagnosis of peripheral pulmonary infectious lesions.* Infection and Drug Resistance, 2020. **13**: p. 567.
211. Shi, C.-L., et al., *Clinical metagenomic sequencing for diagnosis of pulmonary tuberculosis.* Journal of Infection, 2020. **81**(4): p. 567-574.
212. van Boheemen, S., et al., *Retrospective validation of a metagenomic sequencing protocol for combined detection of RNA and DNA viruses using respiratory samples from pediatric patients.* The Journal of Molecular Diagnostics, 2020. **22**(2): p. 196-207.
213. Li, Y., et al., *Application of metagenomic next-generation sequencing for bronchoalveolar lavage diagnostics in critically ill patients.* European Journal of Clinical Microbiology & Infectious Diseases, 2020. **39**(2): p. 369-374.
214. Cantalupo, P.G. and J.M. Pipas, *Detecting viral sequences in NGS data.* Current Opinion in Virology, 2019. **39**: p. 41-48.
215. Fang, X., et al., *Diagnostic value of metagenomic next-generation sequencing for the detection of pathogens in bronchoalveolar lavage fluid in ventilator-associated pneumonia patients.* Frontiers in Microbiology, 2020. **11**: p. 599756.
216. Chen, Y., et al., *Application of metagenomic next-generation sequencing in the diagnosis of pulmonary infectious pathogens from bronchoalveolar lavage samples.* Frontiers in Cellular and Infection Microbiology, 2021. **11**: p. 168.
217. Chen, P., W. Sun, and Y. He, *Comparison of the next-generation sequencing (NGS) technology with culture methods in the diagnosis of bacterial and fungal infections.* Journal of Thoracic Disease, 2020. **12**(9): p. 4924.
218. Wang, Q., et al., *Optimal specimen type for accurate diagnosis of infectious peripheral pulmonary lesions by mNGS.* BMC Pulmonary Medicine, 2020. **20**(1): p. 1-9.
219. Han, D., et al., *mNGS in clinical microbiology laboratories: on the road to maturity.* Critical Reviews in Microbiology, 2019. **45**(5-6): p. 668-685.
220. Wang, H., et al., *Clinical diagnostic application of metagenomic next-generation sequencing in children with severe nonresponding pneumonia.* PLoS ONE, 2020. **15**(6): p. e0232610.
221. Man, W.H., W.A.A. de Steenhuijsen Piters, and D. Bogaert, *The microbiota of the respiratory tract: gatekeeper to respiratory health.* Nature Reviews Microbiology, 2017. **15**(5): p. 259-270.
222. Clinical trial. *Early Identification and Severity Prediction of Acute Respiratory Infectious Disease (ESAR)*. 2021; Available from: https://www.clinicaltrials.gov/ct2/show/record/NCT04955756.

223. Clinical trial. *Pathogen Identification in Pediatric Hematopoietic Stem Cell Transplant Patients With Suspected Lower Respiratory Tract Infection*. 2018; Available from: https://www.clinicaltrials.gov/ct2/show/study/NCT02926612.
224. Clinical trial. *Characterization of Respiratory Microbiota in Susceptibility to Viral Respiratory Infections (RESPIBIOTE)*. 2018; Available from: https://www.clinicaltrials.gov/ct2/show/NCT03600753.
225. trial., C. *mNGS -Guided Antimicrobial Treatment in Early Severe Community-Acquired Pneumonia Among Immunocompromised Patients (MATESHIP)*. 2022; Available from: https://www.clinicaltrials.gov/ct2/show/NCT05290454.
226. High, J., et al., *INHALE: the impact of using FilmArray Pneumonia Panel molecular diagnostics for hospital-acquired and ventilator-associated pneumonia on antimicrobial stewardship and patient outcomes in UK Critical Care—study protocol for a multicentre randomised controlled trial.* Trials, 2021. **22**(1): p. 680.
227. Enne, V.I., et al., *Multicentre evaluation of two multiplex PCR platforms for the rapid microbiological investigation of nosocomial pneumonia in UK ICUs: the INHALE WP1 study.* Thorax, 2022: p. 10.1136/thoraxjnl-2021-216990.
228. Gradisteanu Pircalabioru, G., et al., *Advances in the Rapid Diagnostic of Viral Respiratory Tract Infections.* Front Cell Infect Microbiol, 2022. **12**: p. 807253.
229. Illumina. *Respiratory virus oligo panel*. 2022; Available from: https://www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/respiratory-virus-oligo-panel.html.
230. Paragon Genomics. *CleanPlex® Respiratory Virus Research Panel V2*. 2022; Available from: https://www.paragongenomics.com/product/cleanplex-respiratory-research-panel/
231. vh bio. *Respiratory Virus Targeted Sequencing Panels*. 2022; Available from: https://www.vhbio.com/product/respiratory-virus-targeted-sequencing-panels/
232. Celemics. *Comprehensive respiratory virus panel*. 2022; Available from: https://www.celemics.com/en/sub/products/comprehensive-respiratory-virus-panel.asp
233. Twist Bioscience. *Respiratory Virus Research Panel*. 2022; Available from: https://www.twistbioscience.com/products/ngs/fixed-panels/respiratory-virus-research-panel
234. Phan, M.V.T., et al., *Identification of missed viruses by metagenomic sequencing of clinical respiratory samples from Kenya.* Scientific Reports, 2022. **12**(1): p. 202.
235. Phan, M.V.T., et al., *Unbiased whole-genome deep sequencing of human and porcine stool samples reveals circulation of multiple groups of rotaviruses and a putative zoonotic infection.* Virus Evolution, 2016. **2**(2): p. vew027.
236. Centers for Disease Control and Prevention. *Malaria Biology: Lifecycle*. 2020; Available from: https://www.cdc.gov/malaria/about/biology/index.html.
237. World Health Organization. *Malaria*. 2021; Available from: https://www.who.int/news-room/fact-sheets/detail/malaria.
238. World Health Organization. *World malaria report 2021*. 2021; Available from: https://www.who.int/publications/i/item/9789240040496.
239. World Health Organization. *Diagnostic testing*. 2022; Available from: https://www.who.int/teams/global-malaria-programme/case-management/diagnosis.
240. World Health Organization. *Malaria vaccine: WHO position paper – March 2022*. 2022; Available from: https://www.who.int/publications/i/item/who-wer9709-61%E2%80%9380.
241. World Health Organization. *Statement by the Malaria Policy Advisory Group on the urgent need to address the high prevalence of pfhrp2/3 gene deletions in the Horn of Africa and beyond*. 2021; Available from: https://www.who.int/news/item/28-05-2021-statement-by-the-malaria-policy-advisory-group-on-the-urgent-need-to-address-the-high-prevalence-of-pfhrp2-3-gene-deletions-in-the-horn-of-africa-and-beyond.
242. Landier, J., et al., *Effect of generalised access to early diagnosis and treatment and targeted mass drug administration on Plasmodium falciparum malaria in Eastern Myanmar: an observational study of a regional elimination programme.* The Lancet, 2018. **391**(10133): p. 1916-1926.
243. World Health Organization. *WHO Guidelines for malaria - 31 March 2022*. 2022; Available from: https://apps.who.int/iris/handle/10665/352687.
244. World Health Organization. *Global technical strategy for malaria 2016–2030, 2021 update*. 2021; Available from: https://www.who.int/publications/i/item/9789240031357.
245. Su, X.-z., et al., *Plasmodium Genomics and Genetics: New Insights into Malaria Pathogenesis, Drug Resistance, Epidemiology, and Evolution.* Clinical Microbiology Reviews, 2019. **32**(4): p. e00019-19.
246. Zhou, D., et al., *Genome sequence of Anopheles sinensis provides insight into genetics basis of mosquito competence for malaria parasites.* BMC Genomics, 2014. **15**(1): p. 42.

247. MalariaGEN. *Amplicon sequencing toolkit*. 2022; Available from: https://www.malariagen.net/resources/amplicon-sequencing-toolkit.
248. Oyola, S.O., et al., *Whole genome sequencing of Plasmodium falciparum from dried blood spots using selective whole genome amplification.* Malaria Journal, 2016. **15**(1): p. 597.
249. Neafsey, D.E., A.R. Taylor, and B.L. MacInnis, *Advances and opportunities in malaria population genomics.* Nature Reviews Genetics, 2021. **22**(8): p. 502-517.
250. Hamilton, W.L., et al., *Extreme mutation bias and high AT content in Plasmodium falciparum.* Nucleic Acids Research, 2016. **45**(4): p. 1889-1901.
251. Public Health England. *Malaria reference laboratory: user handbook*. 2021; Available from: https://www.gov.uk/government/publications/malaria-reference-laboratory-mrl-user-handbook.
252. Lee, R.A., et al., *Ultrasensitive CRISPR-based diagnostic for field-applicable detection of Plasmodium species in symptomatic and asymptomatic malaria.* Proceedings of the National Academy of Sciences, 2020. **117**(41): p. 25722-25731.
253. Cunningham, C.H., et al., *A novel CRISPR-based malaria diagnostic capable of Plasmodium detection, species differentiation, and drug-resistance genotyping.* EBioMedicine, 2021. **68**: p. 103415-103415.
254. Nag, S., et al., *Proof of concept: used malaria rapid diagnostic tests applied for parallel sequencing for surveillance of molecular markers of anti-malarial resistance in Bissau, Guinea-Bissau during 2014–2017.* Malaria Journal, 2019. **18**(1): p. 252.
255. Boyce, R.M., et al., *Reuse of malaria rapid diagnostic tests for amplicon deep sequencing to estimate Plasmodium falciparum transmission intensity in western Uganda.* Scientific Reports, 2018. **8**(1): p. 10159-10159.
256. Dalmat, R., et al., *Use cases for genetic epidemiology in malaria elimination.* Malaria Journal, 2019. **18**(1): p. 163.
257. World Health Organization. *Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 8 (2016–2018)*. 2018; Available from: https://apps.who.int/iris/handle/10665/276190.
258. Ashley, E.A., J. Recht, and N.J. White, *Primaquine: the risks and the benefits.* Malaria Journal, 2014. **13**(1): p. 418.
259. Coutrier, F.N., et al., *Laboratory challenges of Plasmodium species identification in Aceh Province, Indonesia, a malaria elimination setting with newly discovered P. knowlesi.* PLoS Neglected Tropical Diseases, 2018. **12**(11): p. e0006924.
260. MalariaGEN. *Where we work*. 2022; Available from: https://www.malariagen.net/about/where-we-work.
261. MalariaGEN. *Parasite*. 2022; Available from: https://www.malariagen.net/parasite.
262. MalariaGEN. *People*. 2022; Available from: https://www.malariagen.net/about/people.
263. MalariaGEN. *GenRe-Mekong*. 2022; Available from: https://www.malariagen.net/parasite/genre-mekong.
264. Ishengoma, D.S., et al., *Deployment and utilization of next-generation sequencing of Plasmodium falciparum to guide anti-malarial drug policy decisions in sub-Saharan Africa: opportunities and challenges.* Malaria Journal, 2019. **18**(1): p. 267-267.
265. *Mobile malaria project*. 2022; Available from: https://mobilemalaria.com/.
266. Runtuwene, L.R., et al., *Nanopore sequencing of drug-resistance-associated genes in malaria parasites, Plasmodium falciparum.* Scientific Reports, 2018. **8**(1): p. 8286.
267. World Health Organization. *Global genomic surveillance strategy for pathogens with pandemic and epidemic potential, 2022–2032*. 2022; Available from: https://www.who.int/publications/i/item/9789240046979.
268. Biosearch Technologies. *ArrayTape*. 2022; Available from: https://www.biosearchtech.com/products/pcr-reagents-kits-and-instruments/consumables/array-tape.
269. Forbes, J.D., et al., *Metagenomics: The Next Culture-Independent Game Changer.* Frontiers in Microbiology, 2017. **8**: p. 1069.
270. Chiu, C.Y. and S.A. Miller, *Clinical metagenomics.* Nature Reviews Genetics, 2019. **20**(6): p. 341-355.
271. Chan Zuckerberg Biohub. *CZ Biohub*. 2022; Available from: https://www.czbiohub.org/infectious-disease/.
272. Rohden, F., et al. *Combined study on Digital Sequence Information (DSI) in public and private databases and traceability*. 2020; Available from: https://www.cbd.int/doc/c/1f8f/d793/57cb114ca40cb6468f479584/dsi-ahteg-2020-01-04-en.pdf.

273. Black, A., et al., *Ten Recommendations for Supporting Open Pathogen Genomic Analysis in Public Health Settings.* Nat Med, 2020. **26**: p. 832–841.
274. Luheshi, L., et al. *Pathogen genomics into practice. PHG Foundation*. 2015; Available from: https://www.phgfoundation.org/report/pathogen-genomics-into-practice.