Health Science Reports

WILEY

Health Science Reports
Open Access

NARRATIVE REVIEW OPEN ACCESS

The Potential of Bioinformatics in Revolutionizing Disease Diagnosis in Sub-Saharan Africa: A Narrative Review

Wilfred Ofosu | Fleischer C. N. Kotey | Alex Odoom | Onyansaniba K. Ntim | Eric S. Donkor

Department of Medical Microbiology, University of Ghana Medical School, Accra, Greater Accra Region, Ghana

Correspondence: Eric S. Donkor (esampane-donkor@ug.edu.gh)

Received: 15 July 2025 | **Revised:** 6 November 2025 | **Accepted:** 18 February 2026

Keywords: bioinformatics | genomics | next-generation sequencing | pathogen detection | personalized medicine | SSA

ABSTRACT

Background and Aim: Despite its role in healthcare delivery and accessibility, the potential of bioinformatics to transform disease diagnosis in SSA remains unknown. This narrative review aims to assess the current status of bioinformatics initiatives in SSA and evaluate the role of bioinformatics in transforming diagnosis of major diseases in the region.

Methods: A narrative review of peer-reviewed literature, policy reports, and program documentation published between 2000 and 2025 was conducted. Sources were identified through targeted database searches and organizational websites (e.g., Africa CDC, H3Africa, genome.gov, etc.). Key themes were synthesized, including bioinformatics pipelines, infrastructure, training and capacity development, integration of sequencing technologies, and disease-specific applications.

Results: Bioinformatics capacity in SSA has expanded primarily through international collaborations and regional networks such as H3Africa and Africa CDC's Pathogen Genomics Initiative. While significant advances in training and expertise are evident, sequencing infrastructure and data management systems remain unevenly distributed. For priority diseases—including malaria, typhoid, HIV, tuberculosis, and neglected tropical diseases—bioinformatics applications show promise in detecting drug resistance, guiding treatment, and enabling genomic surveillance. However, infrastructural barriers such as limited wet-laboratory facilities, data storage capacity, and internet connectivity constrain implementation.

Conclusion: Bioinformatics-based diagnostics have the potential to transform healthcare delivery in SSA by enabling more accurate, timely, and context-specific disease diagnosis. Realizing this potential requires simultaneous investment in sequencing infrastructure, sustainable bioinformatics training, and policies that support data sharing and integration into clinical practice. Strengthening these systems could reduce diagnostic inequities, empower African-led research, and advance the region toward self-reliant and equitable healthcare.

1 | Introduction

Early and accurate disease diagnosis is essential for effective disease treatment, and reducing morbidity and mortality rates associated with delayed diagnosis, incorrect drug prescriptions [1, 2], and continued transmission in hospitalized patients [3]. It is also essential for achieving international health goals, such as the United Nations Sustainable Development Goal 3 (SDG),

Abbreviations: CDC, Center for Disease Control; GATK, Genome Analysis Toolkit; HIV/AIDs, Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome; LMIC, low- or middle-income country; MBPD, multiple bacterial pathogen detection for one health practices; NGS, next-generation sequencing; NTDs, neglected tropical diseases; PCR, polymerase chain reaction; SDG, sustainable development goal; SSA, sub-Saharan Africa; SURPI, sequence-based ultra-rapid pathogen identification; TB, tuberculosis; VEP, variant effect predictor; WGS, whole-genome sequencing.

which aims to prevent deaths of children under 5 years, address neglected tropical diseases (NTDs), and combat infectious diseases like tuberculosis, AIDS, malaria, and hepatitis by 2030. Improved diagnostics, thus reduces the infectious disease burden and addresses emerging pandemic threats on an international scale [4, 5].

The emergence of bioinformatics-based technology has aided in the analysis of the entire human genome sequence, unraveling its intricate structure and deciphering its underlying genetic information [6]. Bioinformatics has also provided a foundation for understanding the complex interplay between genes and environmental factors, evolutionary relationships, and genetic diseases [7–10].

Despite the role of bioinformatics in disease diagnosis [11–13], the adoption and implementation of bioinformatic approaches in healthcare systems in resource-limited settings, is comparatively low to developed countries. Bioinformatics has advanced significantly in developed nations, while gradual progress is being made in developing African countries, with South Africa, Nigeria, and Kenya leading the way [14–16].

The utilization of genomic and bioinformatic approaches in most healthcare settings in SSA is limited [14, 17]. However, this region suffers the highest burden of disease globally; majority of which are infectious [18–21]. Malaria, tuberculosis (TB), typhoid Fever, NTDs, and HIV/AIDS are amongst the major infectious diseases in SSA. Yet, their diagnostic limitations have been reported in several studies, hindering proper disease diagnosis [22–27].

This paper aims to explore the status of bioinformatics initiatives in SSA and evaluate the role of bioinformatics in transforming the diagnosis of major diseases in the region.

2 | Challenges and Prospects for Diagnostic Systems in SSA

Infectious diseases remain the primary cause of chronic illness, premature mortality, and loss of productivity in SSA [21]. The majority of the 122 tests recommended in the List of Essential In Vitro Diagnostics by the World Health Organization (WHO) primarily target infectious diseases, reflecting their global prevalence.

WHO suggests the incorporation of other diagnostic techniques such as nucleic acid amplification tests, automated bench-top analyzers, and microscopy. However, these often require quality control measures, skilled operators, and continuous supply of electricity. Also, many of these tests are either unavailable or rarely accessible [28, 29], or their cost may be high-priced, leading to the utilization of cheaper alternatives that may compromise diagnostic accuracy and reliability [30].

Several studies have highlighted the significant heterogeneity in the accessibility and availability of diagnostic tools across SSA, with socioeconomic and geopolitical factors playing crucial roles [31–34]. Majority of the population can only afford or access a limited range of diagnostic tests provided by primary healthcare facilities or community settings, unlike high-income countries [33].

Most infectious diseases are managed outside well-equipped medical facilities, where access to necessary diagnostic tests is intermittent. As a result, clinicians often rely on syndromic approaches guided by a limited number of diagnostic tests for making diagnoses [35].

Although commonly used in resource-limited settings, syndromic approaches have inherent limitations in terms of specificity and sensitivity. They tend to have poor specificity, leading to unnecessary administration of antimicrobials, while their sensitivity to detect antigens may be low, resulting in missed cases of treatable illnesses [36].

While lateral flow RDTs have shown promise in improving diagnosis in some infectious diseases, their performance can vary significantly. Fischer [37] evaluated the accuracy of RDTs for malaria diagnosis and found considerable variability in sensitivity and specificity among different tests and settings.

The universal accessibility to high-quality laboratory-focused diagnostics in SSA by 2030 has been deemed impractical due to the significant infrastructural upgrades required [2]. Unlike in high-income countries, in SSA, healthcare access and diagnostic proficiency remain a challenge. The exploration of alternative diagnostic strategies holds the potential to elevate clinical care in the region to the level of or even surpass high-income countries.

3 | Diagnostic Limitations of Major Infectious Diseases in SSA

3.1 | Malaria

Microscopic examination is recognized by the Center for Disease Control (CDC) as the gold standard for malaria diagnosis. However, this technique usually requires skilled laboratory experts due to the difficulty in species identification and immature stage identification.

Moreover, the time length of staining and parasite identification during microscopy has delayed diagnosis in symptomatic and asymptomatic patients alike, leading to preventable deaths. Studies have highlighted that RDTs should be complemented with microscopy for optimal diagnostic results [38, 39].

3.2 | Typhoid Fever

Bone Marrow Aspirate Culture is the gold standard for typhoid fever diagnosis owing to its extremely high sensitivity [40]. However, it is invasive and supposedly impracticable in most settings [41].

The use of blood culture is more practical and frequently used. However, it is low in sensitivity (40%–60%) and the production of contamination-free blood culture is difficult in low- or middle-income country (LMIC) settings [28, 29].

3.3 | Human Immunodeficiency Virus (HIV) Infection

Rapid HIV tests yield same-day results and are largely used in HIV testing programs, particularly in settings where trained personnel and laboratory infrastructure are limited.

However, the likeliness false-positives renders results unreliable [30].

WHO advises employing RDTs and serological assays that have a sensitivity of at least 99%, a specificity of at least 98% for the first RDT, 98% for the second, and 99% for the third RDT [42]. However, in the case of children, it is crucial to use nucleic acid testing (NAT) technologies for virological testing to definitively confirm HIV infection [43, 44].

3.4 | Tuberculosis (TB)

Sputum smear microscopy is quick and inexpensive. However, it has limitations in sensitivity [45, 46]. Culture of *M. tuberculosis* is considered the gold standard but is time-consuming.

Molecular diagnostics, such as the GeneXpert MTB/RIF assay, offer significant improvements over traditional methods. This molecular rapid diagnostic tool can simultaneously detect the presence of *M. tuberculosis* and the rifampicin resistance gene, providing more accurate results compared to smear microscopy. However, the molecular approach is costlier and requires sophisticated equipment [47]. WHO strongly recommends the use of the Xpert assay for TB and multidrug-resistant TB diagnosis in HIV-infected individuals [48].

3.5 | Neglected Tropical Diseases (NTDs)

Diagnosis of NTDs typically relies on the microscopic analysis of blood, stool, and urine samples. For instance, intestinal helminths and protozoan infections can be diagnosed by visualizing helminth eggs or larvae, as well as protozoan cysts, oocysts or trophozoites in microscopically prepared stool slides [49, 50].

This requires experienced laboratory professionals for accurate identification of these parasitic elements in the sample. Moreover, the severity of the infection is often correlated with the quantity of the parasitic elements present [51–53].

Molecular techniques are impractical in resource-limited settings [54], while point-of-care diagnostics have improved NTDs diagnosis to a great extent. However, several limitations are bound to their use [55]. In SSA, the lack of rapid, accurate, and user-friendly point-of-care testing methods for NTDs has contributed to their widespread neglect and the underestimation of their impact [54, 56].

4 | Bioinformatics in Human Disease Diagnosis, and Its Advantage Over Conventional Diagnosis

4.1 | Bioinformatics-Based Pathogen Diagnostics

The application of genomic sequencing has transformed the diagnostic landscape by enabling the comprehensive detection of genetic variants, structural abnormalities, and microbial pathogens that underlie human disease. In contrast to traditional diagnostic assays that target one gene or organism at a time, sequencing provides an unbiased, high-resolution view of the genome or metagenome. Coupled with bioinformatics pipelines, sequencing technologies now underpin the identification of rare genetic syndromes, characterization of tumor mutational profiles, surveillance of infectious disease outbreaks, and assessment of microbial dysbiosis.

The first clinically applied sequencing technology, Sanger sequencing, offered unmatched accuracy and read lengths of up to 1 kb, making it the gold standard for many years. Its utility, however, is restricted to small-scale testing due to low throughput and high cost. Today, Sanger sequencing remains important in clinical practice primarily for confirmatory testing of specific variants, but it has been largely superseded by more scalable platforms [57, 58].

Next-generation sequencing (NGS) represents a major advance over conventional Sanger sequencing by enabling massively parallel sequencing of millions of DNA fragments in a single run. Whereas Sanger sequencing is highly accurate but limited to one fragment at a time, NGS delivers whole genomes, exomes, or targeted panels rapidly and at a fraction of the historical cost. This scalability has transformed clinical diagnostics, allowing comprehensive genetic testing that was previously unfeasible.

In practice, NGS makes it possible to identify causal variants in rare genetic disorders, detect somatic mutations in cancer, and uncover infectious pathogens without prior assumptions about their identity. Compared to conventional methods, NGS not only accelerates turnaround times but also increases sensitivity for low-frequency variants and expands the scope of disease profiling to include transcriptomic and epigenomic layers [59]. While Sanger sequencing remains valuable for validating specific variants, NGS has become the cornerstone of modern bioinformatics pipelines in disease diagnosis due to its speed, breadth, and clinical impact [59, 60].

Third-generation sequencing (TGS) technologies, such as Oxford Nanopore sequencing and Pacific Biosciences single-molecule real-time sequencing, extend the capabilities of next-generation platforms by generating long reads directly from single DNA molecules [61]. Unlike short-read approaches that require assembly and can miss complex rearrangements, TGS resolves structural variants, repetitive regions, and haplotype phasing with much greater accuracy. These long-read platforms also detect epigenetic modifications without separate assays, offering additional layers of diagnostic insight.

Clinically, TGS is proving valuable in de novo assembly of pathogenic genomes, identification of structural variants in rare genetic diseases, and real-time surveillance of infectious outbreaks. While historically limited by higher error rates and costs, continuous improvements in chemistry and base-calling algorithms are rapidly enhancing performance [62]. As a result, TGS is increasingly complementing, and in some applications surpassing, earlier methods, particularly where comprehensive structural and epigenetic resolution is essential for diagnosis [61, 62].

Metagenomic sequencing is also of great diagnostic value in infectious diseases. It can detect microbes where cultures fail, and as well detect multiple co-infecting microbes and antimicrobial resistance genes [63]. It is far more sensitive than conventional culture approach of pathogen detection particularly in bronchoalveolar lavage fluid, blood and sputum samples. However, leading challenges include extremely high costs and difficulty of isolating a pathogen signal from the nose of sequences typically generated in a metagenomic approach [64].

Common pipelines for analyzing NGS metagenomic data for pathogen diagnostics include: Sigma [65], PathoScope 2.0 [66], Kraken [67], IDseq [68], DAMIAN [69], PAIPline [70], and SURPI [71]. These pipelines rapidly identify novel and existing pathogens in different diagnostic samples, accurately speciate pathogens to the strain level, and in assembling full-length microbial genomes, which is of great advantage in the investigation of disease outbreak scenarios.

16S sequence analyses as compared to conventional culture methods are very accurate in terms of pathogen detection. It enables identification of pathogens, even those that cannot be cultured or identified phenotypically. Predominantly, they have a higher sensitivity for bacterial species identification [72–74]. However, the accuracy of detecting pathogens through 16S sequencing relies on various factors such as sequencing methods [75], the regions of DNA sequenced [76], and the analytical pipelines applied, including MIP [77], 16SPIP [78], and MBPD [79].

4.2 | Bioinformatics-Based Genetic Diagnosis

Bioinformatics-analyzing-high throughput sequences is the most important in genetic disease diagnosis [80] and has reduced diagnostic times [81]. Moreover, via the investigation of key genes, likely diagnostic biomarkers, key molecular pathways, and drug targets; the sensitivity of diagnosis has improved [82, 83].

Bioinformatic databases enriched with reference genomic data and genetic disease information concerning numerous genetic disorders have been created. These databases collect and curate genetic and clinical data to aid in the interpretation of genetic variants and assist in diagnosing genetic diseases. ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), Human Gene Mutation Database (HGMD) [84], Genome Aggregation Database (gnomAD) [85], DECIPHER [86], Orphanet [87], and PubMed are common databases accessed.

Bioinformatic pipelines for gene disorder diagnosis help automate and streamline the analysis of genetic variants, reducing manual efforts and providing standardized approaches for gene disorder diagnosis. They facilitate the identification of potentially disease-causing variants, prioritize candidate genes, and assist in clinical decision-making for patients with suspected genetic disorders [88, 89]. Some commonly used bioinformatic pipelines for gene disorder diagnosis include: GATK (Genome Analysis Toolkit) [90], VarScan [89], ANNOVAR [91], VEP (Variant Effect Predictor) [88], InterVar [92], and Galaxy [93].

4.3 | Bioinformatics-Based Medical Decision Support

Processed genomic data can be linked to electronic health records and decision-support software (expert systems). Bioinformatics offers the computational analytic methods needed to develop and train the machine learning algorithms that power these decision-support systems [94, 95]. These systems offer valuable insights into patient data, aiding in disease categorization and timely treatment recommendations [96], and can deal with several disease conditions. Though majority can only deal with one disease condition, multi-target systems capable of diagnosing co-infections are available, albeit few [97].

In HIV care, sequencing data analyzed by bioinformatic algorithms can reveal resistance mutations to first-line antiretroviral drugs. These results, once integrated into patient records, guide clinicians in tailoring personalized therapy regimens. Thus, bioinformatics forms the bridge between laboratory sequencing and clinical action.

4.4 | Bioinformatics-Based Health Access Solutions

The emergence of digital technologies driven by bioinformatic algorithmic tools has made healthcare access possible no matter the location. At the comfort of homes, ill patients can access disease information, diagnostic tools, and treatment plans, leading to personalized medicine, where healthcare access, disease diagnosis, and treatment are patient-oriented [98].

Although bioinformatics itself is not telemedicine, cloud-based bioinformatics platforms (e.g., SIMpLE) can lead to efficient disease diagnosis [99–101]. It also allows health workers in resource-limited settings to upload patient data (e.g., genomic sequences and molecular assay results) and receive analyzed diagnostic reports. For instance, programs like veSEQ–HIHV in Zambia used bioinformatics pipelines to automatically generate viral load and resistance profiles from sequencing data, enabling decentralized clinics to access advanced diagnostic information without local high-performance computing.

5 | Status of Bioinformatic Initiatives in SSA

Bioinformatics in SSA has made reasonable progress, with countries like South Africa, Nigeria, Kenya, Ghana, Uganda, and Tanzania making strides in the field [14, 102, 103]. Short courses, workshops, and incorporation of bioinformatic modules in degree programs are increasingly being organized in SSA [14, 15, 102, 104, 105].

Genomic sequencing, a pre-requisite for implementing a bioinformatic analysis is also becoming increasingly available in both diagnostic and research settings [106–109]. High-throughput sequencing technologies are becoming more affordable, making it possible for more African laboratories to use them in clinical and outbreak investigations [110–112]. However, setting up an NGS facility itself can be expensive, requiring US$100,000–US$700,000 in instrument capital alone (before bioinformatics compute and maintenance) [113].

Several continental and national initiatives are steadily building bioinformatics capacity across SSA. Foremost was the NIH-Wellcome-funded H3Africa consortium (2012–2022), which enabled African researchers to lead genomic studies of health and disease. H3Africa invested in infrastructure and training, and by its conclusion had produced over 700 publications. Crucially, it established H3ABioNet, a pan-African bioinformatics network with 28 nodes in 17 countries (16 in Africa) to provide computing resources, data management and training support (genome.gov).

H3Africa also created a biorepository network (in Uganda, Nigeria, and South Africa) to store and share high-quality biospecimens (genome.gov), and it trained hundreds of African students through workshops, courses and degree programs. In parallel, professional groups like the African Society for Bioinformatics & Computational Biology (ASBCB) host conferences and webinars to link researchers across countries. Global partners have joined these efforts: for example, the NIH Fogarty Center's H3Africa Bioinformatics Training Program funds sustainable in-country training centers (fic.nih.gov), and collaborations with EBI/NCBI help adapt international training materials to African needs. Similarly, Africa CDC and the World Health Organization have launched genomics capacity networks—for example, a 2019–2023 effort to set up 12 regional sequencing laboratories—and now run hands-on sequencing and bioinformatics workshops. In fact, in September 2023, the Africa CDC inaugurated its first bioinformatics training course at its new Addis Ababa lab, providing intensive instruction on pathogen genomics analysis. Despite progress in bioinformatics initiatives, the application of bioinformatics-based approaches in routine clinical practice in SSA is low [114, 115].

Significant infrastructural barriers limit bioinformatics implementation in SSA. High-throughput sequencing requires reliable laboratory networks and supply chains; in practice, many African labs still lack in-country next-generation sequencers and must send samples abroad. Even when equipment is available, consumables and maintenance costs can exceed national budgets (devex.com). On the computing side, few institutions have local high-performance clusters or robust data centers, so analysts often rely on limited servers or cloud resources. Connectivity and power issues compound the problem: slow or intermittent internet makes it hard to transfer large genomic datasets or access public databases, and frequent outages disrupt analyses [116]. Data storage infrastructure is underdeveloped—aside from a few centralized repositories (like H3Africa's), most countries lack national bio-data archives. Finally, systematic sample management remains a challenge: beyond the three H3Africa biobanks, there is little coordinated biobanking (cold-chain logistics, standardized metadata, and consented sample collections) across the region (genome.gov). In sum, SSA bioinformatics initiatives have grown rapidly, but "*basic research infrastructure*" must be scaled up in parallel—including sequencing labs, computing/storage systems and biorepositories—to turn training gains into routine genomic research (devex.com) [116].

Generally, implementation of bioinformatics growth in SSA is impeded by several factors. These include: short infrastructural supply and computing systems, limited funding and training opportunities, and lack of expert bioinformatic scientists [14, 117–119]. Very few African institutions operate state-of-the-art genomics labs; acquiring sequencers and reagents is difficult and expensive. Most countries have limited local HPC or storage; inadequate internet bandwidth further constrains data analysis and sharing [116]. Only a handful of biorepositories exist (e.g., H3Africa's three centers genome.gov); many research studies struggle with long-term specimen storage, tracking and distribution. Sustaining infrastructure (power, maintenance, and connectivity) requires ongoing investment. National policies and regulatory frameworks for genomics are still evolving (although H3Africa's ELSI projects have started harmonizing ethics guidelines) [120].

In summary, SSA has made impressive strides through collaborative networks and training programs. The region now boasts a vibrant community of bioinformaticians and a blueprint of successful initiatives (genome.gov). Going forward, however, closing the gap between promise and practice will hinge on parallel investments in hard infrastructure—high-throughput labs, data centers, and biobanks—to fully support the genomics revolution in Africa.

6 | Potential of Bioinformatic-Based Disease Diagnosis in SSA

6.1 | Malaria

High-throughput sequencing of *Plasmodium* parasites is revealing genetic markers of resistance and diagnostic failure. For example, targeted sequencing of the PfKelch13 gene tracks emerging artemisinin resistance, while large whole-genome databases (such as MalariaGEN's Pf7, containing > 20,000 parasite genomes) let researchers detect deletions of the pfhrp2 gene that cause false negatives in common rapid tests. Bioinformatics pipelines analyze these data to map where resistant, or "stealth" parasites are spreading. This genomic surveillance can inform local policies—for instance, indicating which drugs are still effective in a region or whether a given RDT can be trusted (sangerinstitute.blog). By quickly flagging new threats, these methods help health programs adapt in near real-time and deploy resources where they are most needed. Genomic sequences of *Plasmodium* spp. can be also identified using bioinformatic-based pathogen diagnostics databases, including EuPathDB (http://EuPathDB.org; formerly ApiDB).

Despite the potential of bioinformatics in improving malaria diagnosis in the region, the role of RDTs will continue to be significant given their portability, durability, accessibility, and relatively faster turnaround times. Its affordability makes it accessible to diverse populations and settings, consequently bridging the gap between malarial diagnostic availability and accessibility.

6.2 | Typhoid (*Salmonella* Typhi)

Whole-genome sequencing (WGS) and computational analysis are transforming typhoid diagnosis and control. Sequencing *S. Typhi* isolates reveals the specific strain lineages and antibiotic-resistance genes present in a community [121]. In practice, this means that clinicians can use genome data to predict which antibiotics will work against a local outbreak with high accuracy (recent studies show ~99.9% concordance between resistance genes and lab tests) [121]. In SSA, shared bioinformatics platforms (e.g., Africa CDC's SeqAfrica initiative, Pathogenwatch, and TyphiNET dashboards) allow even remote clinics to upload genome data or simplified profiles. These tools aggregate African and global typhoid genomes to show trends in resistance and dominant genotypes [121, 122]. This helps decision-makers choose effective empirical treatments, target vaccination campaigns, and improve water/sanitation measures. By

making analyzed genomic data accessible to all levels of the health system, these approaches extend high-end surveillance into resource-limited settings.

6.3 | Human Immunodeficiency Virus (HIV) Infection

Genomic analysis of HIV can greatly improve diagnosis and care across SSA. By sequencing viral RNA from patients, health systems can track transmission networks and detect drug-resistance mutations in real time [108]. For instance, the veSEQ–HIV platform in Zambia generated full viral genomes from routine blood samples at modest cost. This yielded each patient's viral load and resistance profile simultaneously, enabling clinicians to tailor antiretroviral therapy to the individual's virus [108]. In aggregate, analyzing HIV diversity across regions helps public health officials monitor hot spots and emerging resistant strains. Importantly, as genomic technologies (including portable sequencers and cloud-based analysis) become cheaper and more user-friendly, even clinics with minimal lab infrastructure can benefit [108]. Pilot programs in SSA show that integrating genomics into HIV care can improve outcomes in rural areas and reduce long-term costs. By bringing personalized medicine to the broader population, bioinformatics-driven HIV diagnosis promises more equitable care for people living with HIV in Africa.

6.4 | Tuberculosis

NGS is poised to replace or augment traditional TB tests. While the GeneXpert MTB/RIF assay is often criticized as costly and requiring sophisticated equipment [47], NGS approaches also require substantial investment in wet-lab facilities, sequencing platforms, and sequencing infrastructure.

For instance, GeneXpert MTB/RIF assay typically costs around US$12.90 assay (consumables, reagents, and staff) [123], while setting up an NGS facility can require US$100,000–US$700,000 in instrument capital alone (before bioinformatics compute and maintenance) [113].

In 2023, WHO began recommending targeted NGS (tNGS) for TB diagnosis because it can identify resistance to many drugs in a single test (microbiologyresearch.org). This approach amplifies and sequences multiple TB resistance genes directly from a patient's sputum and uses bioinformatics tools (like TB-Profiler) to call mutations. South Africa and other countries are piloting such pipelines: for example, South Africa has adopted a GeneXpert XDR cartridge and is evaluating full sequencing systems (microbiologyresearch.org). These methods dramatically shorten the time to detect multidrug-resistant TB. Moreover, TB genomics can map transmission chains (who-infected-whom) in communities, helping target interventions. In SSA's context, diagnostic sequencing can be implemented at central or regional labs with results reported electronically to remote clinics. Even without local high-end labs, samples can be sent to network labs and analyzed via online platforms. Overall, genomic TB diagnostics promise faster, more comprehensive testing that aligns with SSA's burden of drug-resistant TB.

6.5 | Neglected Tropical Diseases (NTDs)

For NTDs like schistosomiasis, filariasis, onchocerciasis, and leishmaniasis, bioinformatics is identifying new targets for better tests. Sequencing parasite and vector genomes has uncovered stage-specific proteins and genetic signatures that may serve as sensitive biomarkers [17]. For example, recent proteomic and genomic studies of filarial worms (causing river blindness and lymphatic filariasis) have pinpointed antigens and gene fragments that could improve blood tests or rapid diagnostics [17]. Similarly, comparing genomes of Leishmania or Schistosoma strains can reveal mutations associated with treatment failure. Computational tools can then design multi-pathogen assays (e.g., diagnostic panels or mobile-phone readers) that screen several NTDs at once. EuPathDB (http://EuPathDB.org; formerly ApiDB) contains sequences of NTDs, including Leishmania which can be compared via alignment. In the SSA health context, once these genomic markers are known, they can be incorporated into low-cost tests deployed by field workers. Complex data interpretation is handled by accessible software or cloud services, so community clinics benefit from high-tech development. By enriching elimination programs with genomic insights, these methods help catch persistent infections that evade older tests, bringing cutting-edge tools to underserved populations.

Given the impracticability of PCR in resource-limited settings in relation to NTDs [50], sequencing is even more impractical in these settings. This means that microscopy is still relevant and should be used with strict adherence to laboratory protocols and handled by skilled laboratory personnel.

7 | Conclusion

The field bioinformatics is increasingly developing in SSA, where it has the potential to alleviate the burden imposed by major infectious diseases in the region. By harnessing advances in genomics, computational biology, and data-driven surveillance, the region can move beyond traditional diagnostic limitations toward faster, more accurate, and context-specific solutions. For diseases of major burden, such as malaria, typhoid, HIV, TB, and NTDs, bioinformatics enables early detection of drug resistance, improved test development, and real-time outbreak tracking. These capabilities not only enhance clinical decision-making but also strengthen health systems by bridging gaps between rural and urban settings, where disparities in access to quality diagnostics are most pronounced.

While significant infrastructural barriers remain—such as limited sequencing facilities, insufficient data storage, and unstable connectivity—the growing network of African-led initiatives demonstrates that sustainable progress is achievable. Investment in laboratory infrastructure, biobanks, computational resources, and training will be critical to transforming pilot successes into routine practice. Ultimately, integrating bioinformatics into diagnostic pathways offers more than technological advancement; it represents a path toward equity, resilience, and self-reliance in African healthcare. With continued commitment from governments, regional networks, and international partners, bioinformatics can become a cornerstone of Africa's response to infectious diseases, paving the way for a healthier future.

Author Contributions

Conceptualization: Fleischer C. N. Kotey and Eric S. Donkor. Methodology: Wilfred Ofosu, Fleischer C. N. Kotey, Alex Odoom, Onyansaniba K. Ntim, and Eric S. Donkor. Validation: Wilfred Ofosu, Fleischer C. N. Kotey, Alex Odoom, Onyansaniba K. Ntim, and Eric S. Donkor. Investigation: Wilfred Ofosu, Fleischer C. N. Kotey, Alex Odoom, Onyansaniba K. Ntim, and Eric S. Donkor. Data curation: Wilfred Ofosu, Fleischer C. N. Kotey, Alex Odoom, Onyansaniba K. Ntim, and Eric S. Donkor. Writing – original draft preparation: Wilfred Ofosu, Alex Odoom, Onyansaniba K. Ntim. Writing – review and editing: Wilfred Ofosu, Fleischer C. N. Kotey, Alex Odoom, Onyansaniba K. Ntim, and Eric S. Donkor. Supervision: Eric S. Donkor. Project administration: Eric S. Donkor. Funding acquisition: Eric S. Donkor. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This work was funded by the National Institutes of Health, USA, through the "Application of Data Science to Build Research Capacity in Zoonoses and Food-Borne Infections in West Africa (DS-ZOOFOOD) Training Programme" hosted at the Department of Medical Microbiology, University of Ghana Medical School (Grant Number: UE5TW012566). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Disclosure

The lead author Eric S. Donkor affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Ethics Statement

The authors have nothing to report.

Consent

All authors have agreed to the publication of this review.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors have nothing to report.

References

1. Á. Bocskor, M. Hunyadi, and D. Vince, "National Academies of Sciences, Engineering, and Medicine (2015) The Integration of Immigrants Into American Society. Washington, DC: The National Academies Press," *Intersections* 3, no. 3 (2017): 157, https://doi.org/10.17356/ieejsp.v3i3.401.

2. The PLOS Digital Health Staff The Potential of Digital Molecular Diagnostics for Infectious Diseases in Sub-Saharan Africa," *PLOS Digital Health* 1, no. 6 (2022): e0000064, https://doi.org/10.1371/journal.pdig.0000064.

3. K. E. Kollef, G. E. Schramm, A. R. Wills, R. M. Reichley, S. T. Micek, and M. H. Kollef, "Predictors of 30-Day Mortality and Hospital Costs in Patients With Ventilator-Associated Pneumonia Attributed to Potentially Antibiotic-Resistant Gram-Negative Bacteria," *Chest* 134, no. 2 (2008): 281–287, https://doi.org/10.1378/chest.08-1116.

4. C. D. Kelly-Cirino, J. Nkengasong, H. Kettler, et al., "Importance of Diagnostics in Epidemic and Pandemic Preparedness," *BMJ Global Health* 4, no. Suppl 2 (2019): e001179, https://doi.org/10.1136/bmjgh-2018-001179.

5. M. D. Perkins and P. M. Small, "Partnering for Better Microbial Diagnostics," *Nature Biotechnology* 24, no. 8 (2006): 919–921, https://doi.org/10.1038/nbt0806-919.

6. E. S. Lander, L. M. Linton, B. Birren, et al., "Initial Sequencing and Analysis of the Human Genome," *Nature* 409, no. 6822 (2001): 860–921, https://doi.org/10.1038/35057062.

7. A. Bayat, "Bioinformatics," *BMJ* 324, no. April (2002): 1018–1022, www.ebi.ac.uk.

8. A. Culhane, "Buehler LK, Rashidi HH: Bioinformatics Basics: Applications in Biological Science and Medicine," *BioMedical Engineering OnLine* 5, no. 1 (2006): 41, https://doi.org/10.1186/1475-925x-5-41.

9. J. Pevsner Bioinformatics and Functional Genomics, Second Edition. In; 2009:i-xxviii, https://doi.org/10.1002/9780470451496.fmatter.

10. L. Hood and L. Rowen, "The Human Genome Project: Big Science Transforms Biology and Medicine," *Genome Medicine* 5, no. 9 (2013): 79, https://doi.org/10.1186/gm483.

11. M. Ibañez-Lligoña, S. Colomer-Castell, A. González-Sánchez, et al., "Bioinformatic Tools for NGS-Based Metagenomics to Improve the Clinical Diagnosis of Emerging, Re-Emerging and New Viruses," *Viruses* 15, no. 2 (2023): 587, https://doi.org/10.3390/v15020587.

12. C. Castaneda, K. Nalley, C. Mannion, et al., "Clinical Decision Support Systems for Improving Diagnostic Accuracy and Achieving Precision Medicine," *Journal of Clinical Bioinformatics* 5 (2015): 4, https://doi.org/10.1186/s13336-015-0019-3.

13. L. F. Soualmia and T. Lecroq, "Bioinformatics Methods and Tools to Advance Clinical Care. Findings From the Yearbook 2015 Section on Bioinformatics and Translational Informatics," *Yearbook of Medical Informatics* 10, no. 1 (2015): 170–173, https://doi.org/10.15265/IY-2015-026.

14. T. K. Karikari, "Bioinformatics in Africa: The Rise of Ghana?," *PLoS Computational Biology* 11, no. 9 (2015): e1004308, https://doi.org/10.1371/journal.pcbi.1004308.

15. O. Tastan Bishop, E. F. Adebiyi, A. M. Alzohairy, et al., "Bioinformatics Education—Perspectives and Challenges Out of Africa," *Briefings in Bioinformatics* 16, no. 2 (2015): 355–364, https://doi.org/10.1093/bib/bbu022.

16. O. O. Ojo and M. Omabe, "Incorporating Bioinformatics Into Biological Science Education in Nigeria: Prospects and Challenges," *Infection, Genetics and Evolution* 11, no. 4 (2011): 784–787, https://doi.org/10.1016/j.meegid.2010.11.015.

17. S. Y. Bah, C. M. Morang'a, J. A. Kengne-Ouafo, L. Amenga–Etego, and G. A. Awandare, "Highlights on the Application of Genomics and Bioinformatics in the Fight Against Infectious Diseases: Challenges and Opportunities in Africa," *Frontiers in Genetics* 9 (2018): 575, https://doi.org/10.3389/fgene.2018.00575.

18. K. Mudie, M. M. Jin, L. Tan, et al., "Non-Communicable Diseases in Sub-Saharan Africa: A Scoping Review of Large Cohort Studies," *Journal of Global Health* 9, no. 2 (2019): 020409, https://doi.org/10.7189/jogh.09.020409.

19. M. J. Nyirenda, "Non-Communicable Diseases in Sub-Saharan Africa: Understanding the Drivers of the Epidemic to Inform Intervention Strategies," *International Health* 8, no. 3 (2016): 157–158, https://doi.org/10.1093/inthealth/ihw021.

20. F. Sitas, M. Parkin, Z. Chirenje, et al., "Disease and Mortality in Sub-Saharan Africa," Second Edi. (Jamison DT, Feachem RG, Makgoba

MW, et al., eds.). Washington, D.C. *International Bank for Reconstruction and Development/The World Bank* (2006): https://doi.org/10.1596/978-0-8213-6397-3.

21. World Health Organisation., "A Heavy Burden: The Productivity Cost of Illness in Africa." *Leaving No One Behind in Africa* (2019).

22. O. Ajibola, M. Mshelia, B. Gulumbe, and A. Eze, "Typhoid Fever Diagnosis in Endemic Countries: A Clog in the Wheel of Progress?," *Medicina* 54, no. 2 (2018): 23, https://doi.org/10.3390/medicina54020023.

23. M. Bharadwaj, M. Bengtson, M. Golverdingen, L. Waling, and C. Dekker, "Diagnosing Point-of-Care Diagnostics for Neglected Tropical Diseases," *PLoS Neglected Tropical Diseases* 15, no. 6 (2021): e0009405, https://doi.org/10.1371/journal.pntd.0009405.

24. A. B. Bosco, J. I. Nankabirwa, A. Yeka, et al., "Limitations of Rapid Diagnostic Tests in Malaria Surveys in Areas With Varied Transmission Intensity in Uganda 2017–2019: Implications for Selection and Use of HRP2 RDTs," *PLoS One* 15, no. 12 (2020): e0244457, https://doi.org/10.1371/journal.pone.0244457.

25. M. Hawkes, J. P. Katsuva, and C. K. Masumbuko, "Use and Limitations of Malaria Rapid Diagnostic Testing by Community Health Workers in War-Torn Democratic Republic of Congo," *Malaria Journal* 8 (2009): 308, https://doi.org/10.1186/1475-2875-8-308.

26. N. Masawa, F. Bani, and R. Ndege, "Tuberculosis: Diagnostic Challenges in Rural Africa," *Praxis* 108, no. 15 (2019): 991–996, https://doi.org/10.1024/1661-8157/a003355.

27. N. Ranadive, S. Kunene, S. Darteh, et al., "Limitations of Rapid Diagnostic Testing in Patients With Suspected Malaria: A Diagnostic Accuracy Evaluation From Swaziland, a Low-Endemicity Country Aiming for Malaria Elimination," *Clinical Infectious Diseases* 64, no. 9 (2017): 1221–1227, https://doi.org/10.1093/cid/cix131.

28. V. Mogasale, E. Ramani, V. V. Mogasale, and J. Park, "What Proportion of *Salmonella* Typhi Cases Are Detected by Blood Culture? A Systematic Literature Review," *Annals of Clinical Microbiology and Antimicrobials* 15, no. 1 (2016): 32, https://doi.org/10.1186/s12941-016-0147-z.

29. C. M. Parry, L. Wijedoru, A. Arjyal, and S. Baker, "The Utility of Diagnostic Tests for Enteric Fever in Endemic Locations," *Expert Review of Anti-Infective Therapy* 9, no. 6 (2011): 711–725, https://doi.org/10.1586/eri.11.47.

30. S. Baveewo, M. R. Kamya, H. Mayanja-Kizza, et al., "Potential for False Positive HIV Test Results With the Serial Rapid HIV Testing Algorithm," *BMC Research Notes* 5 (2012): 154, https://doi.org/10.1186/1756-0500-5-154.

31. M. J. Azevedo, "The State of Health System(s) in Africa: Challenges and Opportunities," *Historical Perspectives on the State of Health and Health Systems in Africa, Volume II* 2 (2017): 1–73, https://doi.org/10.1007/978-3-319-32564-4_1.

32. M. Stevenson, "A Situation Analysis of the State of Supply of In Vitro Diagnostics in Low-Income Countries," *Global Public Health* 15, no. 12 (2020): 1836–1846, https://doi.org/10.1080/17441692.2020.1801791.

33. K. A. Fleming, S. Horton, M. L. Wilson, et al., "The Lancet Commission on Diagnostics: Transforming Access to Diagnostics," *Lancet* 398, no. 10315 (2021): 1997–2050, https://doi.org/10.1016/S0140-6736(21)00673-5.

34. K. S. Mooketsane and M. B. Phirinyane, "Health Governance in Sub-Saharan Africa," *Global Social Policy* 15, no. 3 (2015): 345–348, https://doi.org/10.1177/1468018115600123d.

35. M. Urdea, L. A. Penny, S. S. Olmsted, et al., "Requirements for High Impact Diagnostics in the Developing World," *Nature* 444, no. S1 (2006): 73–79, https://doi.org/10.1038/nature05448.

36. C. van Gemert, M. Hellard, C. S. Bradshaw, et al., "Syndromic Management of Sexually Transmissible Infections in Resource-Poor Settings: A Systematic Review With Meta-Analysis of the Abnormal Vaginal Discharge Flowchart for *Neisseria gonorrhoea* and *Chlamydia trachomatis*," *Sexual Health* 15, no. 1 (2018): 1–12, https://doi.org/10.1071/sh17070.

37. P. Fischer, "Faculty Opinions Recommendation of Systematic Review and Meta-Analysis: Rapid Diagnostic Tests Versus Placental Histology, Microscopy and PCR for Malaria in Pregnant Women," *Malaria Journal* 10 (2011): 321, https://doi.org/10.3410/f.13371019.14741134.

38. O. G. Ajakaye and M. R. Ibukunoluwa, "Performance Evaluation of a Popular Malaria RDT in Nigeria Compared With Microscopy," *Journal of Parasitic Diseases* 44, no. 1 (2020): 122–125, https://doi.org/10.1007/s12639-019-01170-y.

39. I. Abdulkadir, H. Rufai, S. Ochapa, et al., "Malaria Rapid Diagnostic Test in Children: The Zamfara, Nigeria Experience," *Nigerian Medical Journal* 56, no. 4 (2015): 278–282, https://doi.org/10.4103/0300-1652.169744.

40. C. L. Kim, L. M. Cruz Espinoza, K. S. Vannice, et al., "The Burden of Typhoid Fever in Sub-Saharan Africa: A Perspective," *Research and Reports in Tropical Medicine* 13 (2022): 1–9, https://doi.org/10.2147/RRTM.S282461.

41. J. Wain and S. Hosoglu, "The Laboratory Diagnosis of Enteric Fever," *Journal of Infection in Developing Countries* 2, no. 06 (2008): 421, https://doi.org/10.3855/jidc.155.

42. World Health Organisation. *WHO Recommendations to Assure HIV Testing Quality*. Geneva; 2015.

43. World Health Organisation. *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach*. 2nd Edition. 2nd Editio. Geneva: World Health Organisation; 2016.

44. World Health Organisation. *WHO Recommendations on the Diagnosis of HIV Infection in Infants and Children*. Switzerland: World Health Organisation; 2010.

45. S. Chakravorty, M. K. Sen, and J. S. Tyagi, "Diagnosis of Extrapulmonary Tuberculosis by Smear, Culture, and PCR Using Universal Sample Processing Technology," *Journal of Clinical Microbiology* 43, no. 9 (2005): 4357–4362, https://doi.org/10.1128/JCM.43.9.4357-4362.2005.

46. K. R. Steingart, V. Ng, M. Henry, et al., "Sputum Processing Methods to Improve the Sensitivity of Smear Microscopy for Tuberculosis: A Systematic Review," *Lancet Infectious Diseases* 6, no. 10 (2006): 664–674, https://doi.org/10.1016/s1473-3099(06)70602-8.

47. D. Arora and B. Dhanashree, "Utility of Smear Microscopy and Genexpert for the Detection of *Mycobacterium tuberculosis* in Clinical Samples," *Germs* 10, no. 2 (2020): 81–87, https://doi.org/10.18683/germs.2020.1188.

48. World Health Organisation. Xpert MTB/RIF Assay for the Diagnosis of Pulmonary and Extrapulmonary TB in Adults and Children. In: Geneva: World Health Organization 2013; 2013, https://apps.who.int/iris/bitstream/handle/10665/112472/9789241506335_eng.pdf.

49. J. Utzinger, S. L. Becker, S. Knopp, et al., "Neglected Tropical Diseases: Diagnosis, Clinical Management, Treatment and Control," *Swiss Medical Weekly* 142 (2012): w13727, https://doi.org/10.4414/smw.2012.13727.

50. L. Meurs, A. M. Polderman, N. V. S. Vinkeles Melchers, et al., "Diagnosing Polyparasitism in a High-Prevalence Setting in Beira, Mozambique: Detection of Intestinal Parasites in Fecal Samples by Microscopy and Real-Time PCR," *PLoS Neglected Tropical Diseases* 11, no. 1 (2017): e0005310, https://doi.org/10.1371/journal.pntd.0005310.

51. L. A. Tchuem Tchuenté, "Control of Soil-Transmitted Helminths in Sub-Saharan Africa: Diagnosis, Drug Efficacy Concerns and Challenges," *Acta Tropica* 120 (2011): S4–S11, https://doi.org/10.1016/j.actatropica.2010.07.001.

52. M. Booth, P. Vounatsou, E. K. N'goran, M. Tanner, and J. Utzinger, "The Influence of Sampling Effort and the Performance of the Kato-Katz Technique in Diagnosing *Schistosoma mansoni* and Hookworm Co-Infections in Rural Cte dIvoire," *Parasitology* 127, no. 6 (2003): 525–531, https://doi.org/10.1017/s0031182003004128.

53. G. Cringoli, L. Rinaldi, M. P. Maurelli, and J. Utzinger, "FLOTAC: New Multivalent Techniques for Qualitative and Quantitative Copromicroscopic Diagnosis of Parasites in Animals and Humans," *Nature Protocols* 5, no. 3 (2010): 503–515, https://doi.org/10.1038/nprot.2009.235.

54. P. J. Hotez and A. Kamath, "Neglected Tropical Diseases in Sub-Saharan Africa: Review of Their Prevalence, Distribution, and Disease Burden," *PLoS Neglected Tropical Diseases* 3, no. 8 (2009): e412, https://doi.org/10.1371/journal.pntd.0000412.

55. S. Campbell and M. L. Landry, "Rapid Microbial Antigen Tests," *Advanced Techniques in Diagnostic Microbiology* 1 (2018): 99–125, https://doi.org/10.1007/978-3-319-33900-9_5.

56. W. Ofosu, *Burden of Enterobiasis in Africa: Overcoming Neglect and Improving Pediatric Health*, eds. Dogan PN (IntechOpen, 2024). Ch. 3, https://doi.org/10.5772/intechopen.1005475.

57. M. B. S. Al-Shuhaib and H. O. Hashim, "Mastering DNA Chromatogram Analysis in Sanger Sequencing for Reliable Clinical Analysis," *Journal of Genetic Engineering and Biotechnology* 21, no. 1 (2023): 115, https://doi.org/10.1186/s43141-023-00587-6.

58. J. Shendure, S. Balasubramanian, G. M. Church, et al., "DNA Sequencing at 40: Past, Present and Future," *Nature* 550, no. 7676 (2017): 345–353, https://doi.org/10.1038/nature24286.

59. H. Satam, K. Joshi, U. Mangrolia, et al., "Next-Generation Sequencing Technology: Current Trends and Advancements," *Biology* 12, no. 7 (2023): 997, https://doi.org/10.3390/biology12070997.

60. E. R. Mardis, "DNA Sequencing Technologies: 2006–2016," *Nature Protocols* 12, no. 2 (2017): 213–218, https://doi.org/10.1038/nprot.2016.182.

61. E. Moustakli, P. Christopoulos, A. Potiris, et al., "Long-Read Sequencing and Structural Variant Detection: Unlocking the Hidden Genome in Rare Genetic Disorders," *Diagnostics* 15, no. 14 (2025): 1803, https://doi.org/10.3390/diagnostics15141803.

62. T. Xiao and W. Zhou, "The Third Generation Sequencing: The Advanced Approach to Genetic Diseases," *Translational Pediatrics* 9, no. 2 (2020): 163–173, https://doi.org/10.21037/tp.2020.03.06.

63. H. Chen, J. Li, S. Yan, et al., "Identification of Pathogen(s) in Infectious Diseases Using Shotgun Metagenomic Sequencing and Conventional Culture: A Comparative Study," *PeerJ (Corta Madera, CA and London)* 9 (2021): e11699, https://doi.org/10.7717/peerj.11699.

64. M. Batool and J. Galloway-Peña, "Clinical Metagenomics—Challenges and Future Prospects," *Frontiers in Microbiology* 14, no. June (2023): 1186424, https://doi.org/10.3389/fmicb.2023.1186424.

65. T. H. Ahn, J. Chai, and C. Pan, "Sigma: Strain-Level Inference of Genomes From Metagenomic Analysis for Biosurveillance," *Bioinformatics* 31, no. 2 (2015): 170–177, https://doi.org/10.1093/bioinformatics/btu641.

66. C. Hong, S. Manimaran, Y. Shen, et al., "PathoScope 2.0: A Complete Computational Framework for Strain Identification in Environmental or Clinical Sequencing Samples," *Microbiome* 2 (2014): 33, https://doi.org/10.1186/2049-2618-2-33.

67. D. E. Wood and S. L. Salzberg, "Kraken: Ultrafast Metagenomic Sequence Classification Using Exact Alignments," *Genome Biology* 15, no. 3 (2014): R46, https://doi.org/10.1186/gb-2014-15-3-r46.

68. K. L. Kalantar, T. Carvalho, C. F. A. de Bourcy, et al., "IDseq-An Open Source Cloud-Based Pipeline and Analysis Service for Metagenomic Pathogen Detection and Monitoring," *GigaScience* 9, no. 10 (2020): giaa111, https://doi.org/10.1093/gigascience/giaa111.

69. M. Alawi, L. Burkhardt, D. Indenbirken, et al., "DAMIAN: An Open Source Bioinformatics Tool for Fast, Systematic and Cohort Based Analysis of Microorganisms in Diagnostic Samples," *Scientific Reports* 9, no. 1 (2019): 16841, https://doi.org/10.1038/s41598-019-52881-4.

70. A. Andrusch, P. W. Dabrowski, J. Klenner, et al., "PAIPline: Pathogen Identification in Metagenomic and Clinical Next Generation Sequencing Samples," *Bioinformatics* 34, no. 17 (2018): i715–i721, https://doi.org/10.1093/bioinformatics/bty595.

71. S. N. Naccache, S. Federman, N. Veeraraghavan, et al., "A Cloud-Compatible Bioinformatics Pipeline for Ultrarapid Pathogen Identification From Next-Generation Sequencing of Clinical Samples," *Genome Research* 24, no. 7 (2014): 1180–1192, https://doi.org/10.1101/gr.171934.113.

72. C. A. Petti, C. R. Polage, and P. Schreckenberger, "The Role of 16S rRNA Gene Sequencing in Identification of Microorganisms Misidentified by Conventional Methods," *Journal of Clinical Microbiology* 43, no. 12 (2005): 6123–6125, https://doi.org/10.1128/JCM.43.12.6123-6125.2005.

73. H. Cai, M. Archambault, and J. F. Prescott, "16S Ribosomal RNA Sequence—Based Identification of Veterinary Clinical Bacteria," *Journal of Veterinary Diagnostic Investigation* 15, no. 5 (2003): 465–469, https://doi.org/10.1177/104063870301500511.

74. A. Manaka, Y. Tokue, and M. Murakami, "Comparison of 16S Ribosomal RNA Gene Sequence Analysis and Conventional Culture in the Environmental Survey of a Hospital," *Journal of Pharmaceutical Health Care and Sciences* 3 (2017): 8, https://doi.org/10.1186/s40780-017-0074-y.

75. R. Winand, B. Bogaerts, S. Hoffman, et al., "Targeting the 16S rRNA Gene for Bacterial Identification in Complex Mixed Samples: Comparative Evaluation of Second (Illumina) and Third (Oxford Nanopore Technologies) Generation Sequencing Technologies," *International Journal of Molecular Sciences* 21, no. 1 (2019): 298, https://doi.org/10.3390/ijms21010298.

76. J. Tremblay, K. Singh, A. Fern, et al., "Primer and Platform Effects on 16S rRNA Tag Sequencing," *Frontiers in Microbiology* 6 (2015): 771, https://doi.org/10.3389/fmicb.2015.00771.

77. Z. Sun, X. Liu, G. Jing, et al., "Comprehensive Understanding to the Public Health Risk of Environmental Microbes Via a Microbiome-Based Index," *Journal of Genetics and Genomics* 49, no. 7 (2022): 685–688, https://doi.org/10.1016/j.jgg.2021.12.011.

78. J. Miao, N. Han, Y. Qiang, T. Zhang, X. Li, and W. Zhang, "16SPIP: A Comprehensive Analysis Pipeline for Rapid Pathogen Detection in Clinical Samples Based on 16S Metagenomic Sequencing," *BMC Bioinformatics* 18, no. Suppl 16 (2017): 568, https://doi.org/10.1186/s12859-017-1975-3.

79. X. Yang, G. Jiang, Y. Zhang, et al., "MBPD: A Multiple Bacterial Pathogen Detection Pipeline for One Health Practices," *iMeta* 2, no. 1 (2023): e82, https://doi.org/10.1002/imt2.82.

80. M. M. S. Al Haggar, "Bioinformatics in High Throughput Sequencing: Application in Evolving Genetic Diseases," *Journal of Data Mining in Genomics & Proteomics* 04, no. 03 (2013): 1000131, https://doi.org/10.4172/2153-0602.1000131.

81. C. J. Saunders, N. A. Miller, S. E. Soden, et al., "Rapid Whole-Genome Sequencing for Genetic Disease Diagnosis in Neonatal Intensive Care Units," *Science Translational Medicine* 4, no. 154 (2012): 154ra135, https://doi.org/10.1126/scitranslmed.3004041.

82. W. Yu, W. Yu, Y. Yang, and Y. Lü, "Exploring the Key Genes and Identification of Potential Diagnosis Biomarkers in Alzheimer's Disease

Using Bioinformatics Analysis," *Frontiers in Aging Neuroscience* 13 (2021): 602781, https://doi.org/10.3389/fnagi.2021.602781.

83. A. Khan, A. Ali, M. Junaid, et al., "Identification of Novel Drug Targets for Diamond-Blackfan Anemia Based on RPS19 Gene Mutation Using Protein-Protein Interaction Network," *BMC Systems Biology* 12, no. Suppl 4 (2018): 39, https://doi.org/10.1186/s12918-018-0563-0.

84. P. D. Stenson, M. Mort, E. V. Ball, et al., "The Human Gene Mutation Database: Towards a Comprehensive Repository of Inherited Mutation Data for Medical Research, Genetic Diagnosis and Next-Generation Sequencing Studies," *Human Genetics* 136, no. 6 (2017): 665–677, https://doi.org/10.1007/s00439-017-1779-6.

85. S. Gudmundsson, M. Singer-Berk, N. A. Watts, et al., "Variant Interpretation Using Population Databases: Lessons From gnomAD," *Human Mutation* 43, no. 8 (2022): 1012–1030, https://doi.org/10.1002/humu.24309.

86. E. Bragin, E. A. Chatzimichali, C. F. Wright, et al., "DECIPHER: Database for the Interpretation of Phenotype-Linked Plausibly Pathogenic Sequence and Copy-Number Variation," *Nucleic Acids Research* 42, no. Database issue (2014): D993–D1000, https://doi.org/10.1093/nar/gkt937.

87. A. Rath, A. Olry, F. Dhombres, M. M. Brandt, B. Urbero, and S. Ayme, "Representation of Rare Diseases in Health Information Systems: The Orphanet Approach to Serve a Wide Range of End Users," *Human Mutation* 33, no. 5 (2012): 803–808, https://doi.org/10.1002/humu.22078.

88. S. E. Hunt, B. Moore, R. M. Amode, et al., "Annotating and Prioritizing Genomic Variants Using the Ensembl Variant Effect Predictor-A Tutorial," *Human Mutation* 43, no. 8 (2022): 986–997, https://doi.org/10.1002/humu.24298.

89. D. C. Koboldt, K. Chen, T. Wylie, et al., "VarScan: Variant Detection in Massively Parallel Sequencing of Individual and Pooled Samples," *Bioinformatics* 25, no. 17 (2009): 2283–2285, https://doi.org/10.1093/bioinformatics/btp373.

90. A. McKenna, M. Hanna, E. Banks, et al., "The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data," *Genome Research* 20, no. 9 (2010): 1297–1303, https://doi.org/10.1101/gr.107524.110.

91. K. Wang, M. Li, and H. Hakonarson, "ANNOVAR: Functional Annotation of Genetic Variants From High-Throughput Sequencing Data," *Nucleic Acids Research* 38, no. 16 (2010): e164, https://doi.org/10.1093/nar/gkq603.

92. Q. Li and K. Wang, "InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines," *American Journal of Human Genetics* 100, no. 2 (2017): 267–280, https://doi.org/10.1016/j.ajhg.2017.01.004.

93. K. Chappell, B. Francou, C. Habib, et al., "Galaxy Is a Suitable Bioinformatics Platform for the Molecular Diagnosis of Human Genetic Disorders Using High-Throughput Sequencing Data Analysis: Five Years of Experience in a Clinical Laboratory," *Clinical Chemistry* 68, no. 2 (2021): 313–321, https://doi.org/10.1093/clinchem/hvab220.

94. K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, "Machine Learning Applications in Cancer Prognosis and Prediction," *Computational and Structural Biotechnology Journal* 13 (2015): 8–17, https://doi.org/10.1016/j.csbj.2014.11.005.

95. H. J. Jang and K. O. Cho, "Applications of Deep Learning for the Analysis of Medical Data," *Archives of Pharmacal Research* 42, no. 6 (2019): 492–504, https://doi.org/10.1007/s12272-019-01162-9.

96. P. S. Patra, D. Sahu, and D. I. Mandal, "An Expert System for Diagnosis of Human Diseases," *International Journal of Computer Applications* 1 (2010): 71, https://doi.org/10.5120/279-439.

97. F. E. Ayo, J. B. Awotunde, R. O. Ogundokun, S. O. Folorunso, and A. O. Adekunle, "A Decision Support System for Multi-Target Disease Diagnosis: A Bioinformatics Approach," *Heliyon* 6, no. 3 (2020): e03657, https://doi.org/10.1016/j.heliyon.2020.e03657.

98. J. P. O. Li, H. Liu, D. S. J. Ting, et al., "Digital Technology, Tele-Medicine and Artificial Intelligence in Ophthalmology: A Global Perspective," *Progress in Retinal and Eye Research* 82 (2021): 100900, https://doi.org/10.1016/j.preteyeres.2020.100900.

99. A. Palumbo, N. Ielpo, B. Calabrese, et al., "SIMpLE: A Mobile Cloud-Based System for Health Monitoring of People With ALS," *Sensors* 21, no. 21 (2021): 7239, https://doi.org/10.3390/s21217239.

100. J. Wang, G. Zhang, W. Wang, K. Zhang, and Y. Sheng, "Cloud-Based Intelligent Self-Diagnosis and Department Recommendation Service Using Chinese Medical BERT," *Journal of Cloud Computing* 10, no. 1 (2021): 4, https://doi.org/10.1186/s13677-020-00218-2.

101. T. Khubone, B. Tlou, and T. P. Mashamba-Thompson, "Electronic Health Information Systems to Improve Disease Diagnosis and Management at Point-of-Care in Low and Middle Income Countries: A Narrative Review," *Diagnostics* 10, no. 5 (2020): 327, https://doi.org/10.3390/diagnostics10050327.

102. E. P. De Villiers and E. Bongcam-Rudloff, "eBioKit Bioinformatics Workshops in Dar es Salaam, Tanzania," *EMBnet.journal.* 20 (2014): 755, https://doi.org/10.14806/ej.20.0.755.

103. R. Naluyange, G. Mboowa, K. Komakech, D. Semugenze, D. P. Kateete, and W. Ssengooba, "High Prevalence of Phenotypic Pyrazinamide Resistance and Its Association With pncA Gene Mutations in *Mycobacterium tuberculosis* Isolates From Uganda," *PLoS One* 15, no. 5 (2020): e0232543, https://doi.org/10.1371/journal.pone.0232543.

104. E. de Villers, "Introductory Course in Bioinformatics in Kenya," *EMBnet.journal.* 15, no. 2 (2009): 29, https://journal.embnet.org/index.php/embnetnews/article/view/13/42.

105. H. Salminen-Mankonen, J. E. Litton, E. Bongcam-Rudloff, et al. The Pan-European Research Infrastructure for Biobanking and Biomolecular Resources: Managing Resources for the Future of Biomedical Research. 2009;15(2):2–40.

106. D. S. Horner, G. Pavesi, T. Castrignano, et al., "Bioinformatics Approaches for Genomics and Post Genomics Applications of Next-Generation Sequencing," *Briefings in Bioinformatics* 11, no. 2 (2009): 181–197, https://doi.org/10.1093/bib/bbp046.

107. S. C. Inzaule, S. K. Tessema, Y. Kebede, A. E. Ogwell Ouma, and J. N. Nkengasong, "Genomic-Informed Pathogen Surveillance in Africa: Opportunities and Challenges," *Lancet Infectious Diseases* 21, no. 9 (2021): e281–e289, https://doi.org/10.1016/S1473-3099(20)30939-7.

108. F. E. Agamah, P. Arbuthnot, P. Tindana, et al., "Advancing the Course of HIV Genomics Research in Africa," *BMJ Global Health* 10 (2025): e019094, https://doi.org/10.1136/bmjgh-2025-019094.

109. B. A. Mensah, N. E. Akyea-Bobi, and A. Ghansah, "Genomic Approaches for Monitoring Transmission Dynamics of Malaria: A Case for Malaria Molecular Surveillance in Sub–Saharan Africa," *Frontiers in Epidemiology* 2 (2022): 939291, https://doi.org/10.3389/fepid.2022.939291.

110. K. A. Johansen Taber, B. D. Dickinson, and M. Wilson, "The Promise and Challenges of Next-Generation Genome Sequencing for Clinical Care," *JAMA Internal Medicine* 174, no. 2 (2014): 275, https://doi.org/10.1001/jamainternmed.2013.12048.

111. H. Candela, R. Casanova-Sáez, and J. L. Micol, "Getting Started in Mapping-by-Sequencing," *Journal of Integrative Plant Biology* 57, no. 7 (2015): 606–612, https://doi.org/10.1111/jipb.12305.

112. M. Teufel and P. Sobetzko, "Reducing Costs for DNA and RNA Sequencing by Sample Pooling Using a Metagenomic Approach," *BMC Genomics* 23, no. 1 (2022): 613, https://doi.org/10.1186/s12864-022-08831-y.

113. M. Helmy, M. Awad, and K. A. Mosa, "Limited Resources of Genome Sequencing in Developing Countries: Challenges and Solutions," *Applied & Translational Genomics* 9 (2016): 15–19, https://doi.org/10.1016/j.atg.2016.03.003.

114. E. Quansah and T. K. Karikari, "Motor Neuron Diseases in Sub-Saharan Africa: The Need for More Population-Based Studies," *BioMed Research International* 2015 (2015): 298409, https://doi.org/10.1155/2015/298409.

115. T. K. Karikari and J. Aleksic, "Neurogenomics: An Opportunity to Integrate Neuroscience, Genomics and Bioinformatics Research in Africa," *Applied & Translational Genomics* 5 (2015): 3–10, https://doi.org/10.1016/j.atg.2015.06.004.

116. E. Agboli, M. Bitew, C. N. Malaka, et al., "Building Pathogen Genomic Sequencing Capacity in Africa: Centre for Epidemic Response and Innovation Fellowship," *Tropical Medicine and Infectious Disease* 10, no. 4 (2025): 90, https://doi.org/10.3390/tropicalmed10040090.

117. S. A. Tishkoff, F. A. Reed, F. R. Friedlaender, et al., "The Genetic Structure and History of Africans and African Americans," *Science* 324, no. 5930 (2009): 1035–1044, https://doi.org/10.1126/science.1172257.

118. E. Matovu, B. Bucheton, J. Chisi, et al., "Enabling the Genomic Revolution in Africa," *Science* 344, no. 6190 (2014): 1346–1348, https://doi.org/10.1126/science.1251546.

119. T. K. Karikari, E. Quansah, and W. M. Y. Mohamed, "Developing Expertise in Bioinformatics for Biomedical Research in Africa," *Applied & Translational Genomics* 6 (2015): 31–34, https://doi.org/10.1016/j.atg.2015.10.002.

120. J. G. Shaffer, F. J. Mather, M. Wele, et al., "Expanding Research Capacity in Sub-Saharan Africa Through Informatics, Bioinformatics, and Data Science Training Programs in Mali," *Frontiers in Genetics* 10 (2019): 331, https://doi.org/10.3389/fgene.2019.00331.

121. Z. A. Dyson, L. Cerdeira, V. Sharma, et al., "The TyphiNET Data Visualisation Dashboard: Unlocking *Salmonella* Typhi Genomics Data to Support Public Health," *Genome Medicine* 17, no. 1 (2025): 51, https://doi.org/10.1186/s13073-025-01470-4.

122. M. E. Carey, Z. A. Dyson, S. Argimón, et al., "Unlocking the Potential of Genomic Data to Inform Typhoid Fever Control Policy: Supportive Resources for Genomic Data Generation, Analysis, and Visualization," *Open Forum Infectious Diseases* 10, no. Suppl_1 (2023): S38–S46, https://doi.org/10.1093/ofid/ofad044.

123. A. W. Kaso and A. Hailu, "Costs and Cost-Effectiveness of Gene Xpert Compared to Smear Microscopy for the Diagnosis of Pulmonary Tuberculosis Using Real-World Data From Arsi Zone, Ethiopia," *PLoS One* 16, no. 10 (2021): e0259056, https://doi.org/10.1371/journal.pone.0259056.