**South African Journal of Information Management**
ISSN: (Online) 1560-683X, (Print) 2078-1865



# Big Data Analytics for Integrated Infectious Disease Surveillance in sub-Saharan Africa

**Authors:**
Mourine S. Achieng[1]
Oluwamayowa O. Ogundaini[1]

**Affiliation:**
[1]Department of Digital Transformation and Innovation, Graduate School of Business Leadership, University of South Africa, Midrand, South Africa

**Corresponding author:**
Mourine Achieng,
achiems@unisa.ac.za

**Dates:**
Received: 23 Feb. 2023
Accepted: 12 July 2023
Published: 30 Jan. 2024

**How to cite this article:**
Achieng, M.S. & Ogundaini, O.O., 2024, 'Big Data Analytics for Integrated Infectious Disease Surveillance in sub-Saharan Africa', *South African Journal of Information Management* 26(1), a1668. https://doi.org/10.4102/sajim.v26i1.1668



**Background:** Infectious disease outbreaks are common in sub-Saharan Africa (SSA). Consequently, integrated public health surveillance has become increasingly essential for the region. Health surveillance systems enable early detection and monitoring of emerging and re-emerging disease outbreaks, thus informing preparedness and response measures. However, complex and intertwined factors obstruct a successful integrated public health surveillance in SSA, with dire consequences.

**Objectives:** The objective of this article was to establish how big data analytics can be used to enhance integrated infectious disease surveillance and response in SSA.

**Method:** The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) was used to identify and select relevant articles. A total of 10 studies that addressed the article's objective were selected.

**Results:** Findings reveal several barriers to the application of big data analytics for public health surveillance in SSA. These include the absence of regulatory and data governance frameworks for big data management in healthcare, disparities in digital health infrastructure across SSA's healthcare systems, and the digital and analytical skills required for data capture and interpretation. The development of regulatory frameworks is essential for the ethical application of analytical technologies such as artificial intelligence.

**Conclusion:** This article's contributions emphasise the need for comprehensive strategies for the application of big data analytics for public health surveillance, as well as addressing barriers to its successful application by highlighting the requirements for an integrated infectious disease surveillance and response system in SSA.

**Contribution:** The article contributes to the body of knowledge on the interplay between the public health space and digital health interventions by emphasising the beneficial applications of big data analytics for surveillance and response to address emerging and re-emerging infectious disease outbreaks in the health systems of sub-Saharan Africa.

**Keywords:** big data analytics; integrated infectious disease surveillance; public health; digital health; health security; health systems; sub-Saharan Africa.

## Introduction

In sub-Saharan Africa (SSA), the public health sector persistently faces numerous and diverse challenges ranging from institutional incapacity, dwindling skilled human resources, data mismanagement, effects of climate change and a prevalence of diseases (Maphumulo & Bhengu 2019; Top, Konca & Sapaz 2020). The prevalence of infectious and non-communicable diseases places a strain on healthcare systems and exacerbates these challenges (Modjadji 2021), to the extent of threatening global health security. As a result, countries in the region are struggling to achieve universal health care coverage, which is a critical component of achieving the United Nations (UN) Sustainable Development Goal (SDG) number 3 – good health and well-being (UN 2015). To achieve this SDG, SSA's healthcare systems must be strengthened so that they can detect, plan and adequately respond to health threats using reliable, accurate, complete and readily available quality data and information (Uwaezuoke 2020). Strengthening healthcare systems would mean that governments must invest in adequate infrastructure and resources to support healthcare-related services (Mremi et al. 2021). The investments serve as a step in the right direction for healthcare systems to reduce the burden of non-communicable and infectious diseases.

Infectious diseases remain the most serious health crisis confronting SSA's healthcare systems (World Health Organization [WHO] 2010; Aborode et al. 2021). Over the last decade, the region has seen devastating outbreaks of highly infectious diseases such as cholera, Lassa fever and yellow fever, as well as Ebola virus disease (EVD) and, more recently, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Nansikombi et al. 2022). Furthermore, SSA is relatively plagued by a high endemic prevalence of human immunodeficiency virus and acquired immunodeficiency syndrome (HIV and AIDS), tuberculosis, malaria and parasitic diseases compared to the Global North. These infectious disease outbreaks necessitate the development of long-term prevention, monitoring, control and treatment strategies by SSA healthcare systems, as well as adequate management of health-related data (Kandel et al. 2020). Infections can spread quickly and cause high morbidity and mortality rates if infectious disease outbreaks are not detected and reported in a timely manner (Mboussou et al. 2019). Consequently, it is critical that SSA healthcare systems implement measures to ensure access to accurate, complete, reliable, relevant and timely health-related data and information. This should allow healthcare systems through public surveillance systems, to respond to and control infectious disease outbreaks while also preventing outbreak spread (Groseclose & Buckeridge 2017).

The WHO defines public health surveillance as 'the continuous, systematic collection, analysis, and interpretation of health-related data needed for planning, implementation, and evaluation of public health practice' (WHO 2017). Surveillance is thus a critical component of infectious disease evaluation, prevention, response and control. However, Aborode et al. (2021) argued that, despite the outbreaks of emerging and re-emerging infectious diseases, insufficient attention has been paid to the strengthening of Africa's health surveillance systems. The fragmentation and lack of interconnectedness of existing electronic and mobile health technologies that enable data management and communication between the three tiers (primary, secondary and tertiary) of healthcare service provision, for example, explains why disease surveillance programmes in several African countries are inadequate (Mahomed et al. 2017). In addition, the lack of readily available health data (Inzaule et al. 2021) and the inconsistent assessment for quality, analysis and use of routine health data for decision-making (Moyo et al. 2022) in several healthcare systems in SSA impede surveillance of infectious diseases.

Inadequate health-related data management in healthcare systems is a constant threat to disease surveillance in SSA. The health sector generates and records a large volume and variety of data that can be analysed to enable adequate infectious disease surveillance. The large volumes and variety of health-related data and information is defined as big data created from paper records and by the adoption of electronic and mobile health technologies that collect patients' information (Akinnagbe, Peiris & Akinloye 2018). Because big data in general is difficult to manage, process and analyse using traditional analytical and statistical techniques, advanced technologies such as artificial intelligence (AI), internet of things (IoT), machine learning and big data analytics tools have become indispensable (Cozzoli et al. 2022). Big data analytics in healthcare entails analysing large volumes of real-time or historical health records from various sources to identify patterns and trends that lead to meaningful insights for detection, predictive and prescriptive decision making (Akinnagbe et al. 2018; Mehta, Pandit & Shukla 2019). During the coronavirus disease 2019 (COVID-19) global pandemic, for example, AI models were used to analyse big data in healthcare (such as patient health history, symptoms, close contacts and proximity) to detect, warn and respond appropriately (Mehta & Shukla 2022). Therefore, advanced data analytics, when implemented properly, can be used to create a knowledge foundation that strengthens the healthcare systems in SSA by enhancing disease management performance and streamlining administrative complexities.

## Research problem

Infectious disease outbreaks in SSA highlight a critical need to strengthen disease surveillance systems for epidemic and pandemic preparedness and response (Aborode et al. 2021; Nkengasong 2020). To achieve the state of preparedness, healthcare systems in SSA and, by extension, Africa must address the underlying issues impeding adequate management of the increasing large volumes of health-related data in its various formats from multiple sources (Aborode et al. 2021; Ayeni, Misra & Omoregbe 2015). A lack of standardised digital health information systems, for example, results in the fragmentation of large volumes of health-related data, erroneous analysis and, ultimately, suboptimal response of surveillance systems to disease outbreaks.

Several studies have identified the causes of Africa's inadequate disease surveillance systems and the urgent need for reforms to strengthen them (Aborode et al. 2021; Mahomed et al. 2017). However, there are limited studies that indicate how healthcare systems in SSA can adequately achieve a state of preparedness and respond to the current and future infectious disease epidemics and pandemics using big data analytical tools (Wolfe et al. 2021).

The unique challenges confronting SSA's healthcare systems have a negative impact on data management and quality, information flow and decision making in response to infectious disease surveillance (Silenou et al. 2020). These effects continue to put a strain on paper records and available infrastructure for digital data-driven medical research, limiting the benefits of evidence-based practice and stifling the interests of healthcare stakeholders. Infectious disease outbreaks spread quickly if they are not detected and reported quickly, resulting in high morbidity and mortality rates and a prolonged recovery.

## Aim and objectives

Public health surveillance systems require adequate health-related data management to mitigate the effects of infectious disease outbreaks. Therefore, accurate, complete, reliable, relevant and timely health information is crucial for early detection, response and control of potential infectious disease outbreaks. This review article seeks to explore the prospective application of big data analytics to support integrated infectious disease surveillance and response in healthcare systems in SSA. The guiding research question is as follows: How can big data analytics enable integrated infectious disease surveillance and response in healthcare systems in SSA?

The objectives of the article are as follows:

- Identify potential benefits and opportunities for adopting big data analytics to enhance integrated infectious disease surveillance and response in SSA's healthcare systems.
- Identify barriers to the successful implementation of big data analytics in SSA's healthcare systems.
- Outline recommendations to tackle challenges and barriers identified in the second objective above.

## Background studies on infectious disease surveillance in sub-Saharan Africa

This section of the article provides a review of the existing literature on infectious disease surveillance and the use of digital health technologies in SSA.

### Infectious disease and surveillance

Infectious diseases are frequently caused by pathogens found in wild animals and insects (Karimuribo et al. 2012; Kavulikirwa & Sikakulya 2022). These pathogens are often contracted, transmitted and spread through human–animal interactions or crop products, wild animal migration and the movement or gathering of people across national borders. It is difficult to limit the potential damage of infectious disease outbreaks to a population and healthcare systems without adequate data on the occurrence, causes, symptoms and extent of spread (Wolfe et al. 2021). As a result, a collaborative One Health (OH) approach is required to establish surveillance mechanisms to collect data, detect, report, respond to and recover from infectious disease outbreaks in humans or animals in a timely manner (Tambo, Adetunde & Olalubi 2018). Real-time or near-real-time surveillance would allow for quick access to diagnostic information on any threat or potential occurrence of an infectious disease outbreak, lowering morbidity and mortality while saving money on disease control measures.

In response to the need for public health surveillance and response systems, the WHO regional committee for Africa developed the Integrated Disease Surveillance and Response (IDSR) strategy between 1998 and 2001 (Fall et al. 2019). The goal of the strategy was to put in place a comprehensive public health surveillance and monitoring system based on evidence to address the rising number of emerging and re-emerging infectious diseases plaguing certain regions of Africa. The IDSR strategy was developed to ensure the efficient flow of health-related data from the three tiers of care service provision in Africa's healthcare systems to the WHO.

The WHO would then be able to make informed decisions in formulating regulatory steps regarding epidemic or pandemic alerts using health-related data from various healthcare systems. This would include managing and distributing necessary resources, as well as issuing travel advisories. However, for a disease surveillance system to be effective, all relevant health-related data must be adequately collected across all the three tiers of healthcare systems using interconnected electronic and mobile health technologies (Dórea & Revie 2021).

Interconnected electronic and mobile health technologies enable health data in various formats, to be collected from various sources and support advance analysis and reporting (Akter & Wamba 2019). The health data indicators collected would include who, where, when, how, symptoms experienced and contacts, resulting in massive amounts of unstructured data. In contrast to paper-based surveillance reporting, big data analytics tools can assist health systems in monitoring, predicting and prescribing how to manage infectious disease outbreaks in real time. Healthcare drones and smartphones equipped with wireless technologies can be used to collect data indicators in difficult-to-reach and remote rural areas for surveillance purposes. Following that, big data analytics can be used to forecast, map, track, monitor and communicate potential epidemics (Cozzoli et al. 2022). However, for big data analytics to be effective in disease surveillance and response, availability of infrastructure such as mobile and telecommunication, transportation and digital networks as well as skilled personnel are critical. In some countries in SSA, challenges such as a lack of digital infrastructure development, social instabilities and poor network coverage hinder adequate health data collection and reporting, which consequently impacts disease surveillance and timely response (Kavulikirwa & Sikakulya 2022).

### Digital technologies and disease surveillance and response in sub-Saharan Africa

Several studies have investigated and assessed the status of digital health technologies application for integrated disease surveillance and response in SSA (Karimuribo et al. 2012; Mremi et al. 2022; Mwabukusi et al. 2014). There is evidence of improved early detection and timely response in some cases, whereas other studies found weak and ineffective surveillance systems to control potential infectious disease outbreaks. The COVID-19 pandemic highlighted the importance of developing resilient disease surveillance and response systems around the world, particularly in Africa, where there were cases of data misrepresentation, underreporting of cases, late response strategies and a lack of readiness to adapt and make informed decisions based on surveillance data (Aborode et al. 2021).

Mwabukusi et al. (2014) conducted a study on the outcomes of a mobile technology-enabled implementation by the Southern African Centre for Infectious Disease Surveillance (SACIDS) to curb infectious diseases in animals. The surveillance system used a client-server model, with mobile devices allowing data to be submitted via SMS and digital forms via the Internet. The aggregated data were stored in a centralised database and downloaded from the server into a Microsoft Excel spreadsheet for analysis and visualisation. The reporting aided in the detection of foot-and-mouth diseases, as well as follow-ups and subsequent quarantine measures for the location, preventing a spread and outbreak that could harm animals and human health. The system's results demonstrated the practical application of digital technology incorporation for surveillance.

However, what the study reveals is that as data aggregation increases over time, it is critical to scale up and use advanced techniques such as big data analytics to predict threats and prescribe appropriate responses.

Mremi et al. (2022) examined the attitudes and practices of members of Tanzania's Council Health Management Teams (CHMT) regarding their capacity to use disease surveillance data in the implementation of a data analysis intervention. From the study, the authors identified that the members' inability to analyse, interpret and report on surveillance data was because of challenges such as incompleteness of routine data during the time of capture and insufficient accessibility of data. Furthermore, there was a tendency to fail to take preventive measures after disease detection in some of the country's healthcare systems. The authors concluded that there is a need to reconsider how infectious disease health-related data is perceived at all levels of healthcare systems. In addition, there is a great need to invest in capacity building across Tanzania's multitiered structure of primary and district health institutions. To reap the benefits of big data analytics, it is necessary to build capacity in the areas of adequate data capture, data analysis, visualisation, interpretation and reporting to make informed decisions and recommend appropriate action plans.

In another study, Mremi et al. (2021) assessed the effectiveness of the IDSR strategy and its implementation outcomes in SSA. Although many countries' surveillance activities have improved in terms of data reporting, the authors reported that in countries such as Nigeria, Malawi and Tanzania, data integration from multiple sources for surveillance is still lacking. Furthermore, in most healthcare systems, the inadequacy of health-related data remains a barrier to evidence-based decision making and action-oriented resolutions. Mremi et al. (2021) noted that big data analytics is underutilised in the implementation of the IDSR strategy in SSA.

In the aftermath of COVID-19, Aborode et al. (2021) emphasised the importance of Africa's disease surveillance system. While the authors admitted that disease surveillance in Africa is inadequate, they highlighted efforts by the region's Centres for Disease Control and Prevention (CDC) to improve response capabilities. Through the Regional Integrated Surveillance and Laboratory Network (RISLNET), a collaborative plan involving key stakeholders was established for pandemic preparedness by equipping laboratories, providing training and establishing a pathway to identify pathogenic transmission quickly. Regional Integrated Surveillance and Laboratory Network was implemented successfully in the eastern and central parts of SSA. Although the authors briefly mentioned electronic health records (EHRs) to assist with data management, they did not mention the potential of advanced tools like big data analytics to improve the quality of response to infectious disease outbreaks.

There are examples in the Global North and Global South of big data analytics being used to develop efficient models that detect and forecast disease occurrences or care management in healthcare (Cichosz, Johansen & Hejlesen 2016; Mehta et al. 2019; Mehta & Shukla 2022). The use of big data analytics can be seen in the identification of early warning signs through screening, the tracking of infected people and the identification of at-risk locations. Souza et al. (2020), for example, created a big data predictive analytics framework for disease analytics. Cichosz et al. (2016) conducted a literature review on the use of predictive models in the screening and management of diabetes complications. Data from social media posts were mined and analysed to identify disease conditions and symptoms to track disease propagation rates and the location of affected people (Asamoah & Sharda 2015). Different sets of data were collected from multiple sources in China to control the rapid spread of COVID-19 in a timely manner (Wu et al. 2020). This section provides lessons to SSA's health systems on how to adapt context-specific technologies and big data analytics to operationalise and achieve the IDSR strategy's goals.

# Research methods and design

A systematic literature review (SLR) is performed in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (Moher, Liberati & Tetzlaff 2009). Systematic literature reviews generally seek to synthesise and summarise existing knowledge rather than generate new knowledge (Okoli & Schabram 2010). Through a thorough and objective synthesis, SLR seeks to answer a specific question rather than provide a general summary of the literature on a topic of interest (Okoli & Schabram 2010). For this review, the following question guided the study: How can big data analytics enable integrated infectious disease surveillance and response in healthcare systems in SSA?

## Search and identification of sources

The search and identification of relevant studies in the literature were done using keywords limited to a 10-year timeframe from 2012 to 2022. The keywords included big data analytics, infectious disease surveillance, response, Africa and sub-Saharan Africa. The 10-year period was considered because of the increased awareness and adoption of digital health technologies in SSA as well as the emergence and re-emergence of infectious diseases such as EVD, Lassa fever, Yellow fever, SARS-COV-2 and Monkeypox, which have resulted in epidemics and a pandemic (Aborode et al. 2021; Moyo et al. 2022; Silenou et al. 2020). As sources for relevant literature, Google Scholar and PubMed were identified and queried. The researchers queried both databases for relevant peer-reviewed publications on the application and use cases of big data analytics to enable integrated infectious disease surveillance in SSA's health systems.

### Procedure for studies selection

The databases were queried using a combination of the following keywords: 'Big data analytics' AND 'infectious disease surveillance' 'response' AND 'Africa' OR 'sub-Saharan Africa'. To expand the search, articles cited in the retrieved relevant publications and those that cited the retrieved articles were considered but were limited by relevance and the period from 2012 to 2022. The queried combination of keywords yielded a total of 117 publications including journal articles, conference proceedings, reports, books, book chapters and other grey literature publications. In total, a search from Google scholar produced a total of 108 publications. Using the same combination of keywords to query the PubMed database, a total of nine publications were retrieved.

### Exclusion and inclusion criteria

The searched studies were selected according to the following inclusion criteria: (1) the study involves at least one SSA country; (2) studies published in English between 2012 and 2022; (3) studies on big data analytics for infectious disease surveillance and (4) studies that had sound scientific and empirical design (qualitative or quantitative or mixed-methods and systematic reviews). Publications written in other languages, studies conducted outside of the SSA context, epidemiology studies on the spread of infectious diseases, non-open access publications and dissertations were all excluded from this study. The article selection, screening and filtering process is illustrated in Figure 1. After applying the exclusion criteria, the remaining publications served as the primary data for this review. Mendeley citation software was used to consolidate and conduct the screening process. Thereafter, the final articles were analysed to identify findings that illustrate application or non-use of big data analytics for integrated infectious disease surveillance in SSA. In addition, the final articles were also analysed to identify the barriers to the successful application of big data analytics for integrated infectious disease surveillance in SSA.

### Ethical considerations

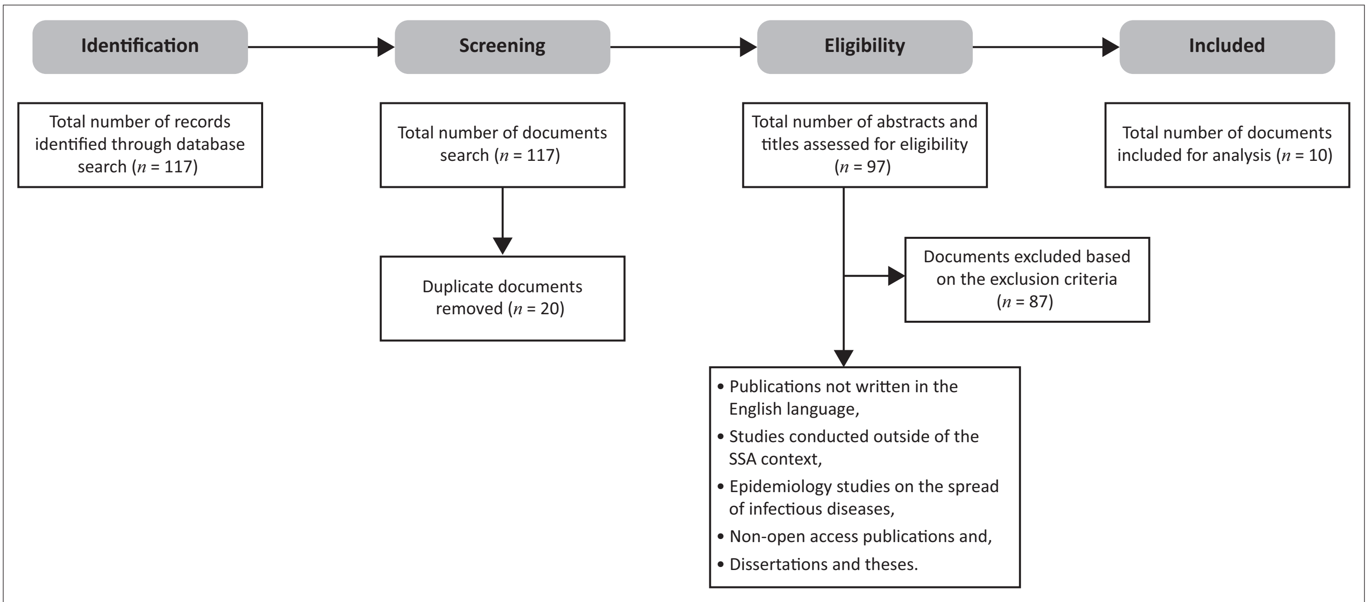
Secondary data were used in this study; therefore, this article does not contain any studies involving human participants performed by any of the authors.

## Results

A total of 117 studies were identified and 20 duplicates were removed. The remaining 97 unique titles and abstracts were assessed for eligibility. After screening the titles and abstracts, 87 publications were excluded based on the exclusion criteria. Finally, 10 studies that addressed big data analytics application in health data management for infectious disease surveillance and response in health systems in SSA were selected for qualitative analysis.

### Study characteristics

Studies on big data and the application of big data analytics to improve infectious disease surveillance and response in health systems in SSA made up the included publications. Out of the 10 publications reviewed, four were country-specific ($n$ = 4), with Tanzania appearing in most of them. Two covered the broader SSA context ($n$ = 2), three of the




SSA, Sub-Saharan Africa.

**FIGURE 1:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of the selection, screening and filtering process.

publications are studies from the West African region of SSA ($n$ = 3), one publication covered the African setting ($n$ = 1). In terms of the infectious diseases, Ebola was the most discussed ($n$ = 4). The remaining publications discussed infectious diseases generally ($n$ = 6). Of the 10 publications reviewed, seven were published in the last five years ($n$ = 7) while the remaining were published between 2014 and 2016 ($n$ = 3).

The upsurge of research outputs on big data analytics application for infectious disease surveillance and response in SSA can be attributed to the rapid digitalisation in the health sector. Based on the included publications, themes such as 'integrated' ($n$ = 3), 'surveillance and response' ($n$ = 4), 'frameworks' ($n$ = 3) emerged. Other emerging themes include 'performance' 'response', 'challenges', 'prospects', 'transmission patterns', 'approaches' and 'strategies'. In analysing the included studies, the adoption and use of big data analytics in health data management for infectious disease surveillance in Africa remains low. However, with the re-emergence and emergence of new infectious diseases such as Ebola, Marburg and COVID-19, there is an enhanced surveillance effort using big data analytical tools and techniques. Table 1 provides the included studies conducted on big data analytics application in health data management to enhance infectious disease surveillance in SSA's healthcare systems.

# Discussion

The SSA region is unique and characterised by contextual realities such as a high likelihood of infectious disease outbreaks, harsh socioeconomic conditions, food insecurity, overburdened public healthcare systems and disparities in infrastructure development. On a positive note, the region has a growing young population, expanding mobile telecoms coverage and massive opportunities for technological innovation and digital transformation. The findings in this review article are interpreted by the authors in terms of the status, barriers and need for big data analytics to adequately

**TABLE 1:** Studies conducted on the use of big data analytics in health data management to enhance infectious disease surveillance in sub-Saharan Africa.

| Title of study | Focus country and Region | Study design or method | Status of big data analytics (mention or actual application) | Opportunities for application for big data analytics of infectious disease surveillance and response |
|---|---|---|---|---|
| Using Big Data Technology to Contain Current and Future occurrence of Ebola Viral Disease and Other Epidemic Disease in West Africa (Ayeni et al. 2015) | West Africa | Mixed methods | Application | **Contribution:** Infectious disease outbreaks can be contained and controlled using big data analytics such as Hadoop MapReduce (HMR).<br>**Opportunities for application for big data analytics:** Securing big data in the healthcare sector. |
| Prospects of big data analytics in Africa healthcare system (Akinnagbe et al. 2018) | Africa | Literature review | Mention | **Opportunities for application for big data analytics:** Analysing big data to fight the HIV epidemic and to develop effective malaria control programmes across sub-Saharan Africa. |
| Big Data Analytics Framework for Childhood Infectious Disease Surveillance and Response System using Modified MapReduce Algorithm (Mwamnyange, Luhanga & Sanket 2021) | Tanzania | Mixed methods | Mention | **Contribution:** A framework that serves as a reference model to make healthcare professionals explore and implement big data analytics technology in the healthcare system.<br>**Opportunities for application for big data analytics:** Intelligence analysis, health data visualisation and decision making. |
| Twenty years of integrated disease surveillance and response in Sub-Saharan Africa: challenges and opportunities for effective management of infectious disease epidemics (Mremi et al. 2021) | SSA | Systematic review | Mention | **Contribution**: *Challenges:* limited skills among health workers because of a lack of training; incomplete and inconsistent data; a lack of use of the generated data. *Opportunities:* effective management of infectious disease epidemics; digital surveillance; event-based surveillance; one health surveillance.<br>**Opportunities for application for big data analytics**: Surveillance systems that are multi-sectoral, multi-disease and multi-indicator epidemic intelligent using health data from health facilities. |
| Recurrent Ebola outbreaks in the eastern Democratic Republic of the Congo: A wake-up call to scale up the integrated disease surveillance and response strategy (Kavulikirwa & Sikakulya 2022) | Democratic Republic of Congo (DRC) | Literature Review | *Mention* | **Opportunities for application for big data analytics:** Provide timely, complete and accurate health data and information. |
| Improving disease surveillance data analysis, interpretation, and use at the district level in Tanzania (Mremi et al. 2022) | Tanzania | Qualitative | *Mention of data management for disease surveillance* | **Contributions:**<br>• Absence of data analysis, interpretation, and utilisation at the sub-national health facilities level.<br>• Inadequate skills and capacity in data management including data analysis, interpretation and use, despite the availability of the guidelines on surveillance data analysis and use.<br>• Disease surveillance performance. |
| Perspectives of different stakeholders on data use and management in public health emergencies in sub-Saharan Africa: a meeting report (Munung et al. 2021) | SSA | Report | *Mention of data management and analysis for disease surveillance* | **Opportunities for application for big data analytics:** Data management and sharing |
| Elucidating transmission patterns from internet reports: Ebola and Middle East respiratory syndrome as case studies (Chowell, Cleaton & Viboud 2016) | Guinea, Sierra Leone, and Liberia | Document analysis | *Mention* | **Opportunities for application for big data analytics:** Enhance detection, forecasting and response to infectious disease threats |
| Characterising Ebola transmission patterns based on internet news reports (Cleaton et al. 2016) | Guinea, Sierra Leone, and Liberia | Document analysis | *Mention of data management and analysis for disease surveillance* | **Opportunities for application for big data analytics:** Provide timely and accurate estimates of key epidemiological parameters. |
| Integrated disease surveillance and response implementation in Liberia, findings from a data quality audit, 2017 (Nagbe et al. 2019) | Liberia | Mixed methods | *Mention of data management and analysis for disease surveillance* | **Opportunities for application for big data analytics:** Decision making purpose and prioritisation in infectious disease surveillance |

SSA, Sub-Saharan Africa.

implement the IDSR strategy in SSA and by extension Africa. This section begins by highlighting the current state of the big data analytics application for integrated infectious disease surveillance and response in SSA.

## Status of big data analytics application for Integrated Disease Surveillance and response in sub-Saharan Africa

There is a scarcity of literature on big data analytics and its applications in the cases of infectious disease surveillance and response in SSA's health systems (Mwamnyange et al. 2021). Based on the literature search and review, there are relatively few studies that mentioned the potential of big data analytics and application cases to realise integrated IDSR strategy in the healthcare systems of SSA (see Table 1). The obvious lack of big data analytics in SSA's healthcare systems for the purpose of infectious diseases surveillance has been attributed to several intricate and intertwined factors (Mahomed et al. 2017; Maphumulo & Bhengu 2019; Top et al. 2020). These factors include inadequate health-related data management, inadequate digital infrastructure and slow rate of adoption of advanced digital health technologies aimed at building resilience in healthcare systems. Many healthcare systems in SSA still use paper-based data collection and record keeping, but in some countries where electronic and mobile health technologies are prevalent, a combination of paper-based and electronic analytics is frequently used (Aborode et al. 2021; Ayeni et al. 2015).

The challenge with paper-based and fragmented analytics in infectious disease surveillance and response is that health-related information is fragmented across healthcare systems (Seitio-Kgokgwe et al. 2016; Tsegaye & Flowerday 2021). These have resulted in some of the SSA health sector's information-related challenges such as duplication, delays in dissemination of relevant and accurate data. The implication of these challenges is the inability to make real-time and actionable health-related decisions on infectious disease outbreaks for surveillance and response. On a positive note, literature indicates that there are efforts and progress made by various stakeholders in Africa's health sector such as the WHO regional committee for Africa to improve on infectious disease surveillance and response (Fall et al. 2019). Currently, the IDSR strategy is being operationalised country by country, which means that each country on the continent has its own independent disease surveillance and response system guided by the strategy. Individual country implementation of the IDSR strategy complicates the collection, analysis and reporting of integrated health-related data for Africa as a whole (Nagbe et al. 2019). The situation presents an opportunity for coordinated collaborations and partnerships.

However, with the increased adoption and use of integrated digital health technologies across SSA's healthcare systems, as well as the presence of the IDSR strategy, there is a promise of coordinated health data management for infectious disease surveillance and response (Akinnagbe et al. 2018; Ayeni et al. 2015; Mremi et al. 2022). The use of digital health technologies enables the collection and integration of large volumes of health data from various sources over time, which can then be integrated and subjected to big data analytics.

Across the continent, there are instances where disease surveillance and response systems are being used for data collection and early detection of potential infectious disease outbreaks (Kavulikirwa & Sikakulya 2022).

For example, Mwabukusi et al. (2014) suggest that surveillance data gathered from animals and humans has improved the detection of potential infectious disease outbreaks. Many SSA healthcare systems, however, use paper-based tools for integrated IDSR. In some cases, Information and Communication Technologies (ICTs) and software applications such as Microsoft Excel are used to collect, analyse, and report on health-related data.

The analysis of big data in the health sector would improve informed decision making process by providing accurate, complete, reliable, relevant and timely health-related data. This can therefore be used in the detecting, forecasting and prescribing imminent infectious disease outbreaks. From the reviewed publications, few studies show how data from the SSA's routine surveillance systems inform action-oriented decisions like quarantines, tracking infected people, alerts and resource allocation to control disease outbreaks, prescribe treatment plans and reduce the rate of transmission (Mremi et al. 2022; Nagbe et al. 2019). For SSA health systems to benefit from the IDSR strategy's outcomes, key healthcare stakeholders such as government, practitioners, researchers and insurers must identify and address barriers to big data analytics application in the SSA context.

## Barriers to big data analytics application for Integrated Disease Surveillance and response systems in sub-Saharan Africa

Several barriers to the successful application of big data analytics for integrated infectious disease surveillance and response in SSA are highlighted across the reviewed literature (Aborode et al. 2021; Akinnagbe et al. 2018; Fall et al. 2019). These barriers are broadly classified in this article into three themes: *inadequate information and communication technology (ICT) and digital infrastructure*, *lack of digital and analytical skills* and *lack of ICT regulatory and data governance frameworks for advanced big data analytics technology application in healthcare.* The extant literature indicates that these barriers have the causal powers to impede using advanced big data analytics tools such as AI and IoTs for public health surveillance. It is therefore important that these barriers are unpacked to gain a better understanding and offer potential solutions if SSA is to achieve an effective integrated infectious disease surveillance and response system.

## Inadequate information and communication technology and digital health infrastructure barrier

The importance of ICT and digital health infrastructure in public health surveillance cannot be overstated.

Because public health surveillance is a continuous systematic collection, analysis and interpretation of data, having an adequate and reliable integrated ICT and digital health infrastructure to support this process is critical. In recent years, Africa has seen improved developments in ICT and digital infrastructure, which has served as a foundation for an increase in digital health initiatives (Hampshire et al. 2021; Holst et al. 2020). However, there have been regional disparities in ICT and digital infrastructure that impede the successful implementation of integrated infectious disease surveillance and response. Information and communication technology and digital infrastructure enable healthcare facilities to collect and share comprehensive, relevant, accurate and timely health-related information. Information sharing and coordinated collaboration are critical for achieving integrated infectious disease surveillance and response in SSA. However, according to Fall et al. (2019), in SSA and by extension Africa, the existing ICT-based surveillance infrastructure is underutilised to support robust detection, response or prevention of infectious diseases in local contexts and at a regional level.

The SSA's healthcare systems are still heavily reliant on paper-based records and reporting (Mwabukusi et al. 2014). Such reporting is prone to human errors, has limited functionality and causes delays that prevent near real-time data analysis in the case of infectious disease surveillance. As a result, where limited technology is used to collect and analyse data, early detection is relatively inefficient. The current state of health surveillance in SSA is such that data are primarily collected from the primary tier of healthcare systems and are often either incomplete or insufficiently disseminated to the higher tiers. The challenge is attributed to the dispersed data collection points, fragmentation of existing healthcare technology, the lack of systems interoperability and standardised reporting (Mremi et al. 2021). As a result, inadequate public health surveillance data impedes adequate analysis and interpretation, making informed decisions and responding to infectious disease outbreaks difficult. Achieving an integrated infectious disease surveillance and response system would involve integrating active and passive health information systems (HISs) into a centralised HIS to collect data on multiple diseases or behaviours of interest to health authorities. For example, country or regional HIS can collect health-related records on multiple infectious diseases and disseminate actionable information to other health systems in different countries.

The impeding contextual realities to technology use in SSA must be addressed for healthcare systems in the region to fully realise an integrated public health surveillance and response. For example, in some areas, unreliable electricity supply is a constant threat to optimal functionality of ICT and digital technologies; therefore, alternate renewable energy sources of electricity need to be considered. Also, cybercrimes, data security and privacy concerns in the health sector are among several factors that deter the use of advanced digital health technologies with big data capabilities. The issues of health data privacy and patient protection necessitate adequate legal frameworks, legislation and cyber-crime prevention strategies in SSA's healthcare systems if the benefits of big data analytics are to be realised (Ayeni et al. 2015).

## Information and communication technology regulatory and data governance frameworks for big data management in healthcare barrier

Despite the analytical benefits and opportunities that advanced technologies such as AI and IoTs bring to public health surveillance, their application in the health sector poses new legal and ethical challenges (Gaffley et al. 2022). Data security and privacy, consent to use, ownership and control of health data, safety, and transparency are examples of challenges emerging from use of advanced technologies particularly in the health sector (Gerke, Minssen & Cohen 2020). An example of the complexities surrounding the use of advanced analytical technologies in public health surveillance is the case of the COVID-19 outbreak, which saw many countries use the contact-tracing application to monitor and prevent the virus's spread. Despite the perceived benefits of the contact-tracing application, it raised privacy and consent to use concerns around the world (Gerke et al. 2020; Gaffley et al. 2022). This example necessitates the establishment of regulatory and data governance frameworks to inform the use of advanced analytical tools particularly in the health sector.

While the majority of the countries in SSA have some form of regulatory and data governance frameworks around health data use, regulatory frameworks that target the use of advanced analytical technologies such as AI are still in their infancy (Akinnagbe et al. 2018; Mandyata, Olowski & Mutale 2017). A lack of a regulatory framework in healthcare setting poses a threat to health data use in an industry with a complex ecosystem. The use of advanced big data analytical technologies necessitates discussions among governments, policymakers and other relevant stakeholders from industry and academia for the purpose of developing regulatory frameworks and policies. This is critical if public health surveillance is to be effective in SSA.

## Inadequate digital and analytical skills

Digital knowledge and skills are critical for detecting, reporting, responding to and recovering from impending infectious disease outbreaks using big data analytics techniques (Mremi et al. 2022). One of the reported weaknesses in surveillance systems in SSA healthcare systems is a lack of knowledge and digital skills to analyse, interpret and report surveillance data. The use of advanced digital health technologies is likely to be resisted because of the need to upskill and re-skill relevant healthcare practitioners to deal with digital health technologies and big data analytics techniques to detect and report infectious disease surveillance data. As a result, personnel training is critical to the application of big data analytics to ensure that surveillance systems in the SSA's healthcare systems are on par with best practices on a global scale.

These three key contextual barriers have negative impact on the realisation of an integrated disease surveillance and response system in SSA and by extension Africa. It is therefore imperative that comprehensive measures are put in place to address these barriers as well as other factors challenging SSA healthcare systems. This is if the operationalisation of the IDSR strategy is to achieve its goal of having a public health surveillance and monitoring system that is based on evidence.

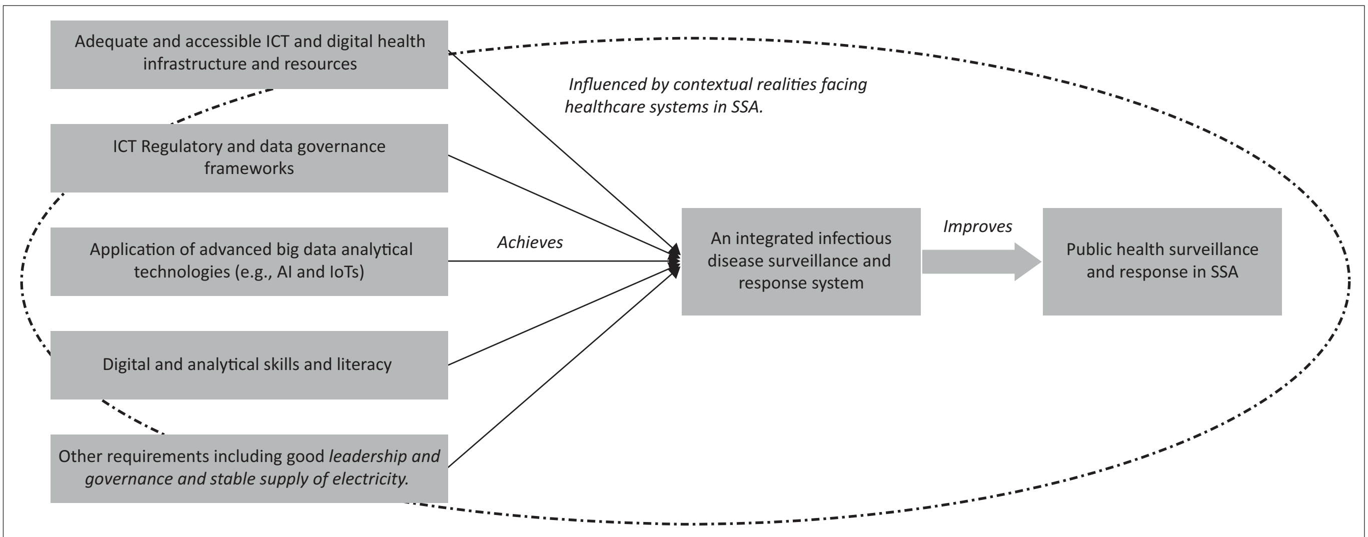
## Requirements for big data analytics application towards an Integrated Disease Surveillance and Response system in sub-Saharan Africa

The use of big data analytics to support integrated public health surveillance of infectious diseases in SSA's healthcare systems necessitates taking proactive steps to address the contextual barriers and challenges identified. It is critical to draw from the barriers identified to clarify the requirements of big data analytics to operationalise the IDSR strategy in SSA if the desired outcomes of preventing, monitoring, controlling and responding to infectious disease outbreaks are to be realised. This study addresses three major requirements that would create an enabling environment for big data analytics application towards an integrated infectious disease surveillance and response system in SSA.

Figure 2 depicts the proposed conceptual framework that highlights the requirements for an integrated infectious disease surveillance and response system. The success of such a system is influenced by the complex and intertwined challenges healthcare systems in SSA face.

Successful application of big data analytics for public health surveillance is heavily reliant on the development of adequate ICT regulatory and data governance frameworks that inform implementation and use of big data technologies such as AI and IoTs in the health sector (Ayeni et al. 2015). The frameworks would guide the operationalisation and expansion of the already existing IDSR strategy, as well as investments in ICT and digital health infrastructure and resources. These investments would equip healthcare workers with the capacity and capability to collect, store and analyse health data using big data analytics techniques, as well as create models to predict infectious disease outbreaks and prescribe appropriate responses. In this way, the shortcomings of public disease surveillance in SSA healthcare systems can be addressed to ensure available ICT resources, data completeness, reduce human errors and improve timely detection and prevention of imminent infectious disease outbreaks reducing high mortality rates (Mwamnyange et al. 2021).

The use of big data analytics techniques in public health surveillance requires specialised digital skills (Mremi et al. 2022). In essence, analytics skills development programmes would ensure that healthcare personnel are upskilled and reskilled to adequately collect and report on disease surveillance data. This is especially important for healthcare professionals who use big data analytical techniques and tools to analyse, interpret and report on disease surveillance data. Popular big data analytics tools include Apache Hadoop, Apache Spark, Microsoft Power BI, data mining, R and Python and natural language processing (NLP) (Mwamnyange et al. 2021; Raguseo 2018). While there is a relative scarcity of skills to use these tools on the continent, it creates an opportunity in the education sectors for the recurriculation of Health Informatics disciplines to develop career professionals such as database administrators, data




ICT, information and communication technology; AI, artificial intelligence; IoTs, internet of things; SSA, sub-Saharan Africa.

**FIGURE 2:** Conceptual framework for an integrated infectious disease surveillance and response system in sub-Saharan Africa.

scientists, data architects and data analysts. As the world increasingly adopts a digital economy, data-driven professions and related stakeholders are likely to influence the budget allocation and ICT infrastructure investments. Personnel with skills and knowledge capacity will make up health leadership required to make informed decisions to deal with emerging and re-emerging disease outbreaks (Ogundaini 2023).

# Conclusions and recommendations

The systematic literature review showed that current disease surveillance systems in Africa's healthcare systems can facilitate early detection and monitoring of infectious diseases, but massive improvements in the aspects of preparedness and responsive action such as prevention, prediction and control are required. There is little evidence of a pressing need to use big data analytics for integrated disease surveillance and response despite the opportunities it presents. As a result, if resilience and health security are to be achieved, African healthcare systems must invest in, adopt and incorporate big data analytics applications capable of surveillance data analysis, interpretation and reporting. Resilience ensures that healthcare systems can withstand the shock, strain and recovery from emerging and re-emerging infectious diseases. The study adds to the knowledge space interplay between public health and digital health by emphasising the potential of big data analytics to address surveillance gaps in Africa's healthcare systems.

Big data analytics requires skills (digital and analytical), but the current reality in SSA is a lack of capacity development and leadership in the areas of digital health (Ogundaini 2023). As a result, it is recommended that data analytics courses be introduced early in health informatics curricula to ensure that home-based community caregivers, social workers and healthcare professionals can collect, detect, report, analyse and interpret data patterns and trends to make informed decisions and action-oriented plans. Capacity development efforts necessitate a multi-stakeholder approach that includes governments developing big data-related policies, academia conducting extensive research and IT practitioners designing training guides and developing user-centred applications tailored to the African context.

The study's limitation is that there may be some relevant publications in non-open access journals that would contribute additional insights to the findings on big data analytics application in Africa's healthcare systems. Because of contextual factors such as limited access to online libraries and paid journals, the authors were unable to access some of the articles. The findings of this study provide an opportunity for researchers, particularly those on the African continent, to expand the study by collecting primary data on infectious disease surveillance from multiple countries. Future research should concentrate on the readiness factors, ethical implications, and strategies that would enable the alignment of big data analytics and the coordination of a timely response to disease outbreaks at the primary and secondary tier levels of public healthcare systems in SSA countries' national health plans.


# Acknowledgements

## Competing interests

The authors declares that they have no financial or personal relationship(s) that may have inappropriately influenced them in writing this article.

## Authors' contributions

M.S.A. and O.O. contributed equally to the conceptualisation and writing of the manuscript. The introduction and method sections were written by M.S.A. while O.O. wrote the background literature section. The results section was jointly written by M.S.A. and O.O. The discussion section was written by M.S.A.

## Funding information

The authors received no financial support for the research, authorship, and/or publication of this article.

## Data availability

Data sharing is not applicable to this article, as no new primary data were created or analysed in this study. The study is a review of openly accessible articles in online journal databases.

## Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.


# References


Aborode, A.T., Hasan, M.M., Jain, S., Okereke, M., Adedeji, O.J., Karra-Aly, A. et al., 2021, 'Impact of poor disease surveillance system on COVID-19 response in africa: Time to rethink and rebuilt', *Clinical epidemiology and global health*, 12(100841), 100841.

Akinnagbe, A., Peiris, K.D.A. & Akinloye, O., 2018, 'Prospects of big data analytics in Africa healthcare system', *Global journal of health science* 10(6), 114.

Akter, S. & Wamba, S.F., 2019, 'Big data and disaster management: A systematic review and agenda for future research', *Annals of Operations Research* 283(1–2), 939–959. https://doi.org/10.1007/s10479-017-2584-2

Asamoah, D. & Sharda, R., 2015, *Adapting CRISP-DM process for social network analytics: Application to healthcare*, *21st Americas Conference on Information Systems (AMCIS 2015)*, Association for Information Systems (AIS), Held 13–15 August 2015, Fajardo.

Ayeni, F., Misra, S. & Omoregbe, N., 2015, 'Using big data technology to contain current and future occurrence of Ebola viral disease and other epidemic diseases in west Africa', in Y. Tan, Y. Shi, F. Buarque, A. Gelbukh, S. Das, A. Engelbrecht (eds.), *Advances in Swarm and Computational Intelligence*, pp. 107–114, Springer International Publishing, Cham.

Chowell, G., Cleaton, J.M. & Viboud, C., 2016, 'Elucidating transmission patterns from Internet reports: Ebola and Middle East respiratory syndrome as case studies', *The Journal of Infectious Diseases* 214(suppl 4), S421–S426. https://doi.org/10.1093/infdis/jiw356

Cichosz, S.L., Johansen, M.D. & Hejlesen, O., 2015, 'Toward big data analytics: Review of predictive models in management of diabetes and its complications: Review of predictive models in management of diabetes and its complications', *Journal of Diabetes Science and Technology* 10(1), 27–34. https://doi.org/10.1177/1932296815611680

Cleaton, J.M., Viboud, C., Simonsen, L., Hurtado, A.M. & Chowell, G., 2016, 'Characterizing Ebola transmission patterns based on Internet news reports', *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 62(1), 24–31. https://doi.org/10.1093/cid/civ748

Cozzoli, N., Salvatore, F.P., Faccilongo, N. & Milone, M., 2022, 'How can big data analytics be used for healthcare organization management? Literary framework and future research from a systematic review', *BMC Health Services Research* 22(1), 809. https://doi.org/10.1186/s12913-022-08167-z

Dórea, F.C. & Revie, C.W., 2021, 'Data-driven surveillance: Effective collection, integration, and interpretation of data to support decision making', *Frontiers in Veterinary Science* 8, 633977. https://doi.org/10.3389/fvets.2021.633977

Fall, I.S., Rajatonirina, S., Yahaya, A.A., Zabulon, Y., Nsubuga, P., Nanyunja, M. et al., 2019, 'Integrated Disease Surveillance and Response (IDSR) strategy: Current status, challenges and perspectives for the future in Africa', *BMJ Global Health* 4(4), e001427. https://doi.org/10.1136/bmjgh-2019-001427

Gaffley, M., Adams, R. & Shyllon, O., 2022, 'Artificial Intelligence. African Insight', A Research Summary of the Ethical and Human Rights Implications of AI in Africa (2022), *HSRC & Meta AI and Ethics Human Rights Research Project for Africa – Synthesis Report*, viewed 20 September 2022, from https://africanaiethics.com/wp-content/uploads/2022/02/Artificial-Intelligence-African-Insight-Report.pdf.

Gerke, S., Minssen, T. & Cohen, G., 2020, 'Ethical and legal challenges of artificial intelligence-driven healthcare', in A. Bohr & K. Memarzadeh (eds.), *Artificial Intelligence in Healthcare*, pp. 295–336, Academic Press.

Groseclose, S.L. & Buckeridge, D.L., 2017, 'Public health surveillance systems: Recent advances in their use and evaluation', *Annual Review of Public Health* 38(1), 57–79. https://doi.org/10.1146/annurev-publhealth-031816-044348

Hampshire, K., Mwase-Vuma, T., Alemu, K., Abane, A., Munthali, A., Awoke, T. et al., 2021, 'Informal mhealth at scale in Africa: Opportunities and challenges', *World development* 140(105257), 105257. https://doi.org/10.1016/j.worlddev.2020.105257

Holst, C., Sukums, F., Radovanovic, D., Ngowi, B., Noll, J. & Winkler, A.S., 2020, 'Sub-Saharan Africa-the new breeding ground for global digital health', *The Lancet. Digital health* 2(4), e160–e162. https://doi.org/10.1016/S2589-7500(20)30027-3

Inzaule, S.C., Tessema, S.K., Kebede, Y., Ogwell Ouma, A.E. & Nkengasong, J.N., 2021, 'Genomic-informed pathogen surveillance in Africa: Opportunities and challenges', *The Lancet Infectious Diseases* 21(9), e281–e289. https://doi.org/10.1016/S1473-3099(20)30939-7

Kandel, N., Chungong, S., Omaar, A. & Xing, J., 2020, 'Health security capacities in the context of COVID-19 outbreak: An analysis of International Health Regulations annual report data from 182 countries', *Lancet* 395(10229), 1047–1053. https://doi.org/10.1016/S0140-6736(20)30553-5

Karimuribo, E.D., Sayalel, K., Beda, E., Short, N., Wambura, P., Mboera, L.G. et al., 2012, 'Towards one health disease surveillance: The Southern African Centre for infectious disease surveillance approach', *The Onderstepoort Journal of Veterinary Research* 79(2), 454. https://doi.org/10.4102/ojvr.v79i2.454

Kavulikirwa, O.K. & Sikakulya, F.K., 2022, 'Recurrent Ebola outbreaks in the eastern Democratic Republic of the Congo: A wake-up call to scale up the integrated disease surveillance and response strategy', *One health (Amsterdam, Netherlands)* 14, 100379. https://doi.org/10.1016/j.onehlt.2022.100379

Mahomed, S., Mahomed, O., Sturm, A.W., Knight, S. & Moodley, P., 2017, 'Challenges with surveillance of healthcare-associated infections in intensive care units in South Africa', *Critical Care Research and Practice* 2017, 1–7. https://doi.org/10.1155/2017/7296317

Mandyata, C.B., Olowski, L.K. & Mutale, W., 2017, 'Challenges of implementing the integrated disease surveillance and response strategy in Zambia: A health worker perspective', *BMC Public Health* 17(1), 746. https://doi.org/10.1186/s12889-017-4791-9

Maphumulo, W.T. & Bhengu, B.R., 2019, 'Challenges of quality improvement in the healthcare of South Africa post-apartheid: A critical review', *Curationis* 42(1), e1–e9. https://doi.org/10.4102/curationis.v42i1.1901

Mboussou, F., Ndumbi, P., Ngom, R., Kamassali, Z., Ogundiran, O., Van Beek, J. et al., 2019, 'Infectious disease outbreaks in the African region: Overview of events reported to the World Health Organization in 2018', *Epidemiology and Infection* 147(e299), e299. https://doi.org/10.1017/S0950268819001912

Mehta, N., Pandit, A. & Shukla, S., 2019, 'Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study', *Journal of Biomedical Informatics* 100(103311), 103311. https://doi.org/10.1016/j.jbi.2019.103311

Mehta, N. & Shukla, S., 2022, 'Pandemic analytics: How countries are leveraging big data analytics and artificial intelligence to fight COVID-19?,' *SN Computer Science* 3(1), 54. https://doi.org/10.1007/s42979-021-00923-y

Modjadji, P., 2021, 'Communicable and non-communicable diseases coexisting in South Africa', *The Lancet. Global Health* 9(7), e889–e890. https://doi.org/10.1016/S2214-109X(21)00271-0

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. & PRISMA Group, 2009, 'Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement', *Annals of Internal Medicine* 151(4), 264–269, W64. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Moyo, E., Musuka, G., Murewanhema, G., Moyo, P. & Dzinamarira, T., 2022, 'Monkeypox outbreak: A perspective on Africa's diagnostic and containment capacity', *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases* 123, 127–130. https://doi.org/10.1016/j.ijid.2022.08.016

Mremi, I.R., George, J., Rumisha, S.F., Sindato, C., Kimera, S.I. & Mboera, L.E.G., 2021, 'Twenty years of integrated disease surveillance and response in Sub-Saharan Africa: Challenges and opportunities for effective management of infectious disease epidemics', *One Health Outlook* 3(1), 22. https://doi.org/10.1186/s42522-021-00052-9

Mremi, I.R., Sindato, C., Kishamawe, C., Rumisha, S.F., Kimera, S.I. & Mboera, L.E.G., 2022, 'Improving disease surveillance data analysis, interpretation, and use at the district level in Tanzania', *Global Health Action* 15(1), 2090100. https://doi.org/10.1080/16549716.2022.2090100

Munung, N.S., Chi, P.C., Abayomi, A., Afolabi, M.O., Ambe, J., Bonarwolo, K. et al., 2021, 'Perspectives of different stakeholders on data use and management in public health emergencies in sub-Saharan Africa: A meeting report'*Wellcome Open Research* 6, 11. https://doi.org/10.12688/wellcomeopenres.16494.2

Mwabukusi, M., Karimuribo, E.D., Rweyemamu, M.M. & Beda, E., 2014, 'Mobile technologies for disease surveillance in humans and animals', *The Onderstepoort Journal of Veterinary Research* 81(2), E1–E5. https://doi.org/10.4102/ojvr.v81i2.737

Mwamnyange, M., Luhanga, E. & Sanket, 2021, 'Big data analytics framework for childhood infectious disease surveillance and response system using modified MapReduce algorithm', *International Journal of Advanced Computer Science and Applications: IJACSA* 12(3). https://doi.org/10.14569/IJACSA.2021.0120345

Nagbe, T., Yealue, K., Yeabah, T., Rude, J.M., Fallah, M., Skrip, L. et al., 2019, 'Integrated disease surveillance and response implementation in Liberia, findings from a data quality audit, 2017', *The Pan African Medical Journal* 33(Suppl 2), 10. https://doi.org/10.11604/pamj.supp.2019.33.2.17608

Nansikombi, H.T., Kwesiga, B., Arinaitwe, E.S., Kwagonza, L. & Harris, J.R., 2022, *Weekly surveillance data reporting on epidemic prone disease, Uganda*, Uganda National Institute of Public Health, viewed 20th September 2022, from https://uniph.go.ug/weekly-surveillance-data-reporting-on-epidemic-prone-diseases-uganda-2020-2021/?utm_source=rss&utm_medium=rss&utm_campaign=weekly-surveillance-data-reporting-on-epidemic-prone-diseases-uganda-2020-2021

Nkengasong, J., 2020, 'Let Africa into the market for COVID-19 diagnostics', *Nature* 580(7805), 565. https://doi.org/10.1038/d41586-020-01265-0

Ogundaini, O.O., 2023, 'eHealth Leadership in Sub-Saharan Africa: A Rapid Review', *2023 IST-Africa Conference*, Tshwane, South Africa, May 31, 2023, pp. 1–8. https://doi.org/10.23919/IST-Africa60249.2023.10187775.

Okoli, C. & Schabram, K., 2010, 'A guide to conducting a systematic literature review of information systems research', *SSRN Electronic Journal* 10(26). https://doi.org/10.2139/ssrn.1954824

Raguseo, E., 2018, 'Big data technologies: An empirical investigation on their adoption, benefits and risks for companies', *International Journal of Information Management* 38(1), 187–195.

Seitio-Kgokgwe, O., Mashalla, Y., Seloilwe, E. & Chida, N., 2016, *Utilization of the District Health Information Software (DHIS) in Botswana: From paper to electronic based system*, *2016 IST-Africa Week Conference*, IEEE, Durban, South Africa, May 11–13, 2016, pp. 1–10. https://doi.org/10.1109/ISTAFRICA.2016.7530690.

Silenou, B.C., Tom-Aba, D., Adeoye, O., Arinze, C.C., Oyiri, F., Suleman, A.K. et al., 2020, 'Use of surveillance outbreak response management and analysis system for human Monkeypox outbreak, Nigeria, 2017–2019', *Emerging Infectious Diseases* 26(2), 345–349. https://doi.org/10.3201/eid2602.191139

Souza, J., Leung, C.K. & Cuzzocrea, A., 2020, 'An innovative big data predictive analytics framework over hybrid big data sources with an application for disease analytics', in L. Barolli, F. Amato, F. Moscato, T. Enokido & M. Takizawa (eds.), *Advanced Information Networking and Applications*, AINA 2020, vol 1151, Advances in Intelligent Systems and Computing, Springer, Cham.

Tambo, E., Adetunde, O.T. & Olalubi, O.A., 2018, 'Re-emerging Lassa fever outbreaks in Nigeria: Re-enforcing "One Health" community surveillance and emergency response practice', *Infectious Diseases of Poverty* 7(1), 37. https://doi.org/10.1186/s40249-018-0421-8

Top, M., Konca, M. & Sapaz, B., 2020, 'Technical efficiency of healthcare systems in African countries: An application based on data envelopment analysis', *Health policy and Technology* 9(1), 62–68. https://doi.org/10.1016/j.hlpt.2019.11.010

Tsegaye, T. & Flowerday, S., 2021, 'A system architecture for ensuring interoperability in a South African national electronic health record system', *South African Computer Journal* 33(1), 1–32. https://doi.org/10.18489/sacj.v33i1.838

UN General Assembly, 2015, A/RES/70/1 transforming our World: The 2030 agenda for sustainable development, UN General Assembly, New York, NY.

Uwaezuoke, S.N., 2020, 'Strengthening health systems in Africa: The COVID-19 pandemic fallout', *Journal of the Pan African Thoracic Society* 1(15), 15–19. https://doi.org/10.25259/JPATS_14_2020

World Health Organization Regional Office for Africa, 2010, *Technical guidelines for integrated disease surveillance and response in the African region*, viewed 18 October 2022, from https://www.afro.who.int/publications/technical-guidelines-integrated-disease-surveillance-and-response-african-region-0.

World Health Organization Regional Office for Africa, 2017, *WHO | Public health surveillance*, viewed 18 October 2022, from https://www.who.int/topics/public_health_surveillance/en/.

Wolfe, C.M., Hamblion, E.L., Dzotsi, E.K., Mboussou, F., Eckerle, I., Flahault, A. et al., 2021, 'Systematic review of Integrated Disease Surveillance and Response (IDSR) implementation in the African region', *PLoS One* 16(2), e0245457. https://doi.org/10.1371/journal.pone.0245457

Wu, J., Wang, J., Nicholas, S., Maitland, E. & Fan, Q., 2020, 'Application of big data technology for COVID-19 prevention and control in China: Lessons and recommendations', *Journal of Medical Internet Research* 22(10), e21980. https://doi.org/10.2196/21980