

## Modern Epidemiological Approaches to Disease Surveillance and Outbreak Predictions

Muhammad Arif<sup>1</sup>

<sup>1</sup> Deputy District Health Officer (Deputy DHO), Health Department Sohbatpur, Baluchistan,  
Email: [dr.arifdajli@gmail.com](mailto:dr.arifdajli@gmail.com)

***DOI: <https://doi.org/10.63163/jpehss.v4i1.1004>***

### Abstract

This paper explores the evolution of epidemiological approaches to disease surveillance and outbreak prediction, transitioning from traditional indicator-based systems to advanced data-driven methodologies. It examines the integration of big data, artificial intelligence (AI), machine learning (ML), natural language processing (NLP), digital epidemiology, genomic surveillance, wastewater-based epidemiology (WBE), and advanced modeling techniques like agent-based models (ABM) and spatial-temporal hybrids. Key discussions include the role of non-traditional data streams such as social media, mobile data, wearables, and IoT in enabling real-time monitoring and early warning systems. Regional challenges in low- and middle-income countries (LMICs), particularly Pakistan and Balochistan, are highlighted, alongside case studies from Ebola outbreaks in Africa. Ethical considerations, data privacy, and future directions toward global pathogen networks and equitable health systems are also addressed. The paper emphasizes a multidisciplinary One Health framework to enhance global health security and prevent pandemics.

**Keywords:** Epidemiological Surveillance, Outbreak Prediction, Artificial Intelligence, Machine Learning, Digital Epidemiology, Genomic Sequencing, Wastewater-Based Epidemiology, Agent-Based Modeling, Global Health Security, One Health

### Introduction

The systematic approach to detecting, verifying, and analyzing public health threats, historically referred to as epidemic intelligence, has transitioned into a highly sophisticated digital discipline (Mayaki, 2025). Traditional surveillance frameworks primarily relied on manual analysis of structured healthcare data, often flowing from official public health reports through paper-based or rudimentary electronic systems (Cazap et al., 2016). These indicator-based systems, while foundational, are increasingly recognized as insufficient in a globalized era characterized by the rapid emergence of infectious diseases. Conventional methods frequently manifest significant delays in reporting and suffer from substantial coverage gaps, particularly in regions where healthcare infrastructure is fragmented or underdeveloped (Abihiro et al., 2014). The contemporary landscape of infectious disease surveillance (IDS) is defined by a shift toward the integration of extensive data streams often categorized as big data to support real-time monitoring and predictive modeling (Cureus, 2025). The World Health Organization (WHO) has advocated for Integrated Disease Surveillance and Response (IDSR), a strategy designed to strengthen national monitoring by coordinating surveillance, laboratory, and response functions into a single cohesive network (WHO, 2024). This evolution is driven by the growing frequency of emerging pathogens, exacerbated by globalization, rapid urbanization, climate change, and the persistent threat of

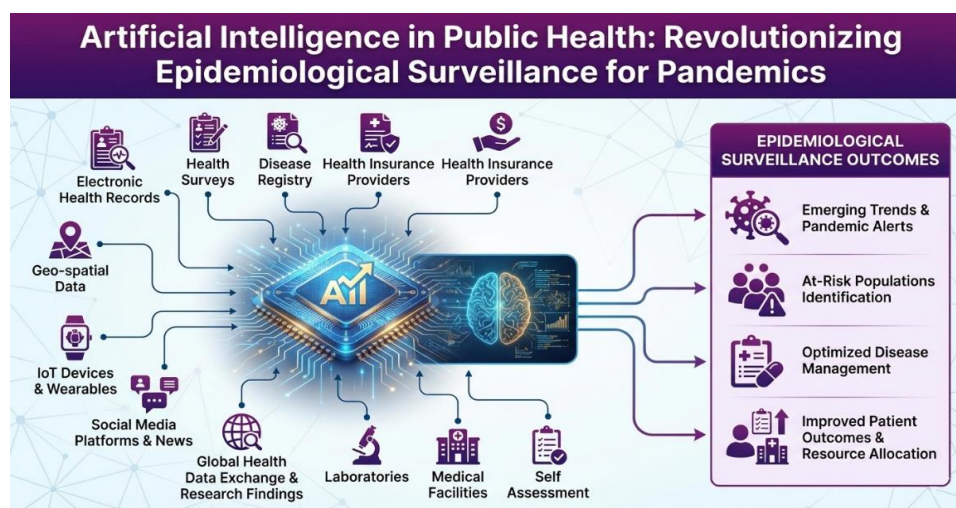
antimicrobial resistance (Salam et al., 2023). Modern approaches leverage the "five Vs" of big data: volume, velocity, variety, veracity, and value. By synthesizing information from diverse sources, including electronic medical records, social media, search engine trends, and environmental sensors, public health agencies can generate contextual insights that were previously unattainable (Aijaz et al., 2023). For instance, the ability to correlate signals from non-healthcare platforms allows systems to identify an escalating outbreak through independent sources confirming rising cases of unknown respiratory illnesses well before formal clinical confirmation (Tzavellas et al., 2026).

**Table 1. Comparison of Traditional Indicator-Based Surveillance and Modern Data-Driven Digital Surveillance Systems**

Surveillance Component	Traditional Approach (Indicator-Based)	Modern Approach (Data-Driven/Digital)
<b>Data Source</b>	Structured official health reports (Mayaki, 2025)	Heterogeneous big data (Web, social media, IoT) (Cureus, 2025)
<b>Reporting Speed</b>	Weekly or monthly delayed reporting (Mayaki, 2025)	Real-time or near real-time data flow (Mayaki, 2025)
<b>Data Type</b>	Primarily clinical/laboratory confirmed (Cureus, 2025)	Syndromic, behavioral, and environmental signals (Cureus, 2025)
<b>Analysis Method</b>	Manual review and basic statistics (Mayaki, 2025)	AI, Machine Learning, and Deep Learning (Mayaki, 2025)
<b>Outreach</b>	Centralized/Hospital-focused (Pakistan Journal of Medical Research, 2024)	Decentralized, including community-based and WBE (Cureus, 2025)

### Artificial Intelligence and Machine Learning in Outbreak Prediction

The integration of artificial intelligence (AI) into public health infrastructure represents a transformative leap in early warning systems (EWS). AI technologies process large, complex datasets at unprecedented speeds, enabling the identification of subtle patterns and anomalies that might indicate the emergence of a disease outbreak (Villanueva-Miranda et al., 2025). The use of machine learning (ML), deep learning (DL), and natural language processing (NLP) has become prevalent, often integrating epidemiological data with climate and wastewater-based indicators (Rony et al., 2023).



**Figure 1. Artificial Intelligence in Public Health Prediction**

### **Supervised Learning and Classification Architectures**

Supervised learning models are fundamental to modern outbreak detection, as they learn from labeled historical data to predict disease presence or transmission risk. Common algorithms such as Random Forests, Support Vector Machines (SVM), and gradient boosting classifiers have demonstrated superior performance in identifying early cases and distinguishing true infectious signals from background environmental noise. These models excel in classifying symptoms and estimating the severity of outbreaks by analyzing high-dimensional patient data (ResearchGate, 2024). Random Forest architectures are particularly valued for their ensemble approach, which utilizes multiple decision trees to handle noisy and complex epidemiological datasets. Meanwhile, SVMs are effective at handling both linear and non-linear relationships, identifying optimal hyperplanes to separate data classes for precision epidemiology (Zinzuwadia et al., 2024). The application of these models extends to identifying high-risk populations and tailoring interventions to specific geographic or demographic communities (Gado et al., 2020).

### **Deep Learning and Time-Series Forecasting**

Advanced deep learning architectures, including recurrent neural networks (RNNs) and transformer-based models, are increasingly employed to estimate outbreak trajectories (Gado et al., 2020). These models are uniquely suited for temporal data, capturing long-term dependencies in disease transmission that classical statistical methods might overlook. A notable development in this area is the use of Graph Transformer Neural Networks to predict multi-regional spread by incorporating human mobility patterns (Wu et al., 2020). By treating different regions as nodes in a graph and modeling movement between them, systems like the Metapopulation Graph Transformer (M-Graphormer) can estimate complex disease parameters and predict trends even when local data is sparse or of low quality (Mienye et al., 2024). This shift from single-region prediction to multi-regional spatial-temporal modeling allows public health authorities to anticipate how an outbreak might cross borders, enabling more effective social distancing or travel restriction strategies (PMC11717196, 2024).

### **The Role of Natural Language Processing (NLP)**

Natural Language Processing serves as the bridge between unstructured digital information and actionable epidemiological intelligence. Systems like BlueDot, HealthMap, and GPHIN utilize NLP to scan thousands of global news sources, airline ticketing data, and social media feeds in real-time. These systems identify clusters of unusual disease activity often days or weeks before official reporting (Cureus, 2025). NLP-driven sentiment analysis and symptom monitoring on social platforms provide behavioral cues that precede clinical diagnosis. For example, during the COVID-19 pandemic, monitoring keywords related to respiratory distress and public concern allowed researchers to track the spread of the virus and the impact of public health messaging (Vishal, 2025). Furthermore, AI-powered chatbots leveraging NLP handle real-time health queries and disseminate prevention knowledge, effectively acting as an interactive layer of the surveillance system (Agbaakin et al., 2024).

### **Digital Epidemiology and Non-Traditional Data Streams**

The field of digital epidemiology leverages "digital exhaust" data generated by everyday interactions with technology to complement traditional clinical surveillance. This approach is particularly valuable in regions where formal healthcare access is limited or where there is a high prevalence of asymptomatic transmission (Croce et al., 2023).

### **Social Media and Search Engines**

Platforms such as X (formerly Twitter) and Google Search Trends have established themselves as vital "event-based" surveillance tools. By analyzing the frequency of symptom-specific search terms, systems can predict influenza activity with high temporal resolution (Panagiotou et al., 2016). While early attempts like Google Flu Trends (GFT) eventually overshot actual case numbers due to changes in user behavior, modern hybrid models like DeepFlu combine neural networks with classical statistical methods to maintain accuracy over time (Barros et al., 2020).

### **Mobile Data and Geospatial Intelligence**

Human mobility is a primary driver of large-scale infectious disease outbreaks. Geographic Information Systems (GIS) and mobile phone tracking provide geospatial insights into how populations move between regions, allowing for the identification of transmission foci (Lessani et al., 2024). Satellite imagery further enhances this by monitoring environmental conditions, such as stagnant water or changes in land use, that contribute to the proliferation of vector-borne diseases like malaria and Zika. Integrating this geospatial data with population density allows for the creation of heat maps that guide the allocation of mobile testing and vaccination units (CDC, 2022).

### **Wearable Technologies and the Internet of Things (IoT)**

The proliferation of smartwatches and physiological sensors has introduced a continuous stream of real-time health data. Devices that monitor heart rate, respiratory rate, body temperature, and blood oxygen levels can identify population-level physiological anomalies that precede the awareness of symptoms by the individuals themselves (Dias, 2020). By 2026, the focus of digital health is expected to shift toward integrating these consumer-grade devices into official clinical workflows, moving from episodic check-ins to continuous health insight (S3 Connected Health, 2024).

### **Genomic Surveillance and Pathogen Characterization**

Genomic sequencing has become an essential tool for understanding the evolution and transmission of infectious pathogens. By decoding the genetic information of viruses and bacteria, epidemiologists can identify specific mutations that might alter transmissibility, virulence, or resistance to treatments (Giovanetti, 2023).

### **Next-Generation Sequencing (NGS) and Metagenomics**

Next-generation sequencing technologies allow for the rapid identification of both known and novel pathogens. Unlike traditional qPCR, which requires specific primers and can only detect a targeted organism, metagenomic NGS can sequence all genetic material in a sample, enabling the discovery of emerging threats (Sutcliffe et al., 2024). Portable sequencing platforms, such as those provided by Oxford Nanopore Technologies (ONT), have enabled real-time surveillance in field conditions, significantly reducing the time required for sample transport and analysis (Bhutta et al., 2021).

### **Phylogenetic Analysis and Variant Tracking**

During the 2018–2020 Ebola outbreak in the Democratic Republic of Congo (DRC), genomic surveillance represented approximately 24% of laboratory-confirmed cases. Phylogenetic analysis of these sequences allowed researchers to detect the emergence of the A82V variant, which was associated with increased human infectivity (Mbala-Kingebeni et al., 2021). These insights directly informed public health decision-making by revealing hidden transmission chains and identifying specific "hot" health zones that required intensive intervention (ITG Research, 2024).

**Table 2. Characteristics and Advantages of Targeted and Untargeted Sequencing Methods in Genomic Surveillance**

Sequencing Approach	Description	Primary Advantage
<b>Amplicon Sequencing</b>	Targeted amplification of viral genome regions	High sensitivity for lineage classification and variant detection (PMC11608212, 2024)
<b>Hybrid Capture</b>	Uses probes to bind specific genetic regions	Captures whole genomes for detailed phylogenetic analysis (PMC11608212, 2024)
<b>WGS (Whole Genome)</b>	Comprehensive decoding of the entire pathogen genome	Definitive identification of mutations and evolutionary tracking (CDC, 2022)
<b>Metagenomics</b>	Sequencing all DNA/RNA in a sample	Discovery of novel viruses and pathogens (Cureus, 2025)

**Wastewater-Based Epidemiology (WBE)**

Wastewater-based epidemiology has emerged as a transformative, non-invasive method for population-level health assessment. Pathogens shed in the stool and urine of infected individuals including those who are asymptomatic can be detected in municipal sewage systems (Zhu et al., 2025).

**Methodological Framework and Sampling**

The reliability of WBE depends on rigorous sampling strategies that account for the spatial and temporal variability of wastewater systems. One significant challenge is the degradation of viral RNA during transport through sewage networks. Advanced stabilization techniques, such as the use of porous superabsorbent polymer (PSAP) beads, have been developed to improve recovery efficiency and the accuracy of concentration detection (Sutcliffe et al., 2024). WBE serves as an early warning signal, often detecting changes in infection trends 4 to 6 days before those same changes are reflected in clinical case reports (CDC, 2022). During the COVID-19 pandemic, the CDC's National Wastewater Surveillance System (NWSS) expanded to include more than 1,200 testing sites across the United States, providing a critical layer of situational awareness that was independent of individual testing behaviors (Lizot et al., 2023).

**Integration with Genomic Tools**

The integration of NGS with WBE allows for the identification of multiple pathogen lineages within a single community sample. Tools like COJAC and Freyja deconvolve the complex genetic mixtures found in wastewater, enabling authorities to track the relative proportions of different variants circulating in the population (Farkas., et al.,2025). This approach is particularly effective for monitoring antimicrobial resistance (AMR) and foodborne pathogens, as it captures a broader representative sample of the community than clinical testing alone (Psomopoulos et al., 2025).

**Advanced Modeling and Simulation Techniques**

The intersection of computer science and epidemiology has given rise to sophisticated models that simulate the spread of disease through human populations with high granularity (Gordeev et al., 2025)

**Agent-Based Models (ABM)**

Agent-based models represent a paradigm shift from population-level equations to individual-level simulations. In these models, "agents" representing individuals follow social, movement, and interaction rules. This allows for the emergence of complex epidemic patterns based on social heterogeneity and adaptive, goal-driven behavior (Becker et al., 2024). In a typical ABM, the

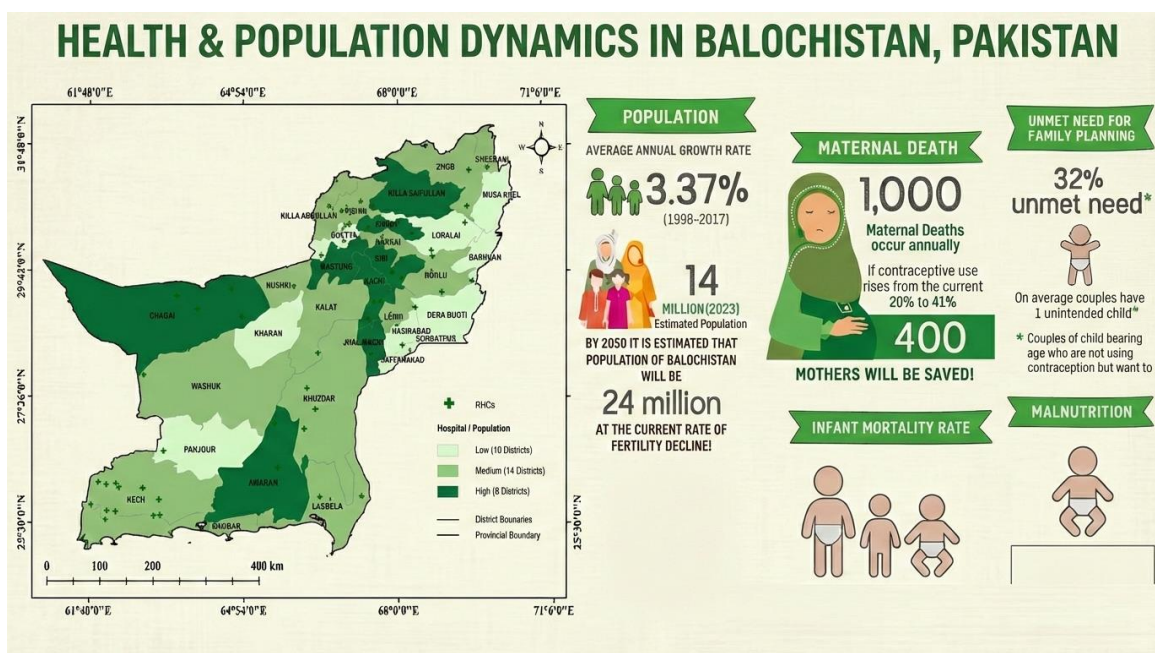
transmission probability ( $\pi_i$ ) for a susceptible agent ( $i$ ) is calculated based on the number of infectious agents ( $k$ ) within a contact distance ( $\theta$ ). The formula  $\pi_i = 1 - (1 - \rho)^k$  represents this, where  $\rho$  is the base probability of infection per contact (Kim et al., 2022). These models are invaluable for evaluating the impact of specific interventions, such as school closures or localized quarantine measures, in a virtual environment before implementation (Amona et al., 2025).

### Spatial-Temporal Hybrid Models

While ABMs provide granular detail, they are computationally intensive. Hybrid models have been developed to balance this granularity with efficiency by spatially coupling discrete agent-based populations with continuous compartmental models, such as Ordinary Differential Equations (Zhang et al., 2021). This allows for detailed analysis of disease transmission in high-density or high-risk areas while simulating the rest of the population more broadly. These models are essential for understanding how movement patterns between diverse geographic regions affect the peak time and shape of the infectious curve (Lakmali et al., 2024).

### Regional Challenges and Implementation: Pakistan and Balochistan

The implementation of integrated surveillance systems faces unique hurdles in low- and middle-income countries (LMICs), where infrastructure gaps and socioeconomic factors influence data quality and response effectiveness (Bashir et al., 2023).



**Figure 2: Demographic Dynamics and Healthcare Infrastructure Challenges in Balochistan**

### The IDSR Journey in Pakistan

Pakistan's Integrated Disease Surveillance and Response System (IDSRS) has evolved from a basic disease registry into a tool for tracking trends across 158 districts (Ali et al., 2025). Despite this progress, significant challenges remain, including a lack of coordination between vertical health programs and a heavy reliance on public sector data, which excludes the 70% of the population that utilizes private healthcare (Pakistan Journal of Medical Research, 2024).

**Table 3. Regional Disease Surveillance Reporting Compliance and Predominant Diseases in Pakistan (Week 26, 2025)**

Region	Expected Reports (Week 26, 2025)	Compliance Rate	Top Reported Diseases
<b>Sindh</b>	2903	97%	Malaria, Acute Diarrhea, ILI (NIH, 2025b)
<b>AJK</b>	404	94%	Acute Diarrhea, ILI, TB (NIH, 2025b)
<b>Gilgit Baltistan</b>	405	93%	VPDs, Mumps, Brucellosis (NIH, 2025b)
<b>Khyber Pakhtunkhwa</b>	2320	64%	Malaria, Measles, Brucellosis (NIH, 2025b)
<b>Balochistan</b>	1307	43%	Malaria, Acute Diarrhea, Pertussis (NIH, 2025b)

In Balochistan, the largest and least developed province of Pakistan, surveillance is particularly hampered by geographic vastness and corruption that has led to a significant number of "ghost" healthcare facilities. A WHO investigation found that while records claimed 1,661 facilities, 257 were nonexistent on the ground, and many others lacked basic amenities like water, sanitation, and tool sterilization (Balochistan Post, 2017).

#### **Localized Gaps in Panjgur District**

Panjgur district in Balochistan exemplifies the struggle for effective surveillance in remote areas. During mid-2025, Panjgur's reporting compliance was as low as 18%, with only 7 out of 38 expected reports received by the national system (Raisani et al., 2025). The disease burden in the area is dominated by Acute Diarrhea, Malaria, and Influenza-like Illnesses, reflecting gaps in clean water access and vector control. The Balochistan Health Institutions Reforms Act of 2025 aims to address these issues by upgrading facilities like the Teaching DHQ Hospital Panjgur to provide better diagnostic and healthcare services (Government of Balochistan, 2025).

#### **Localized Gaps in Sohbatpur District**

Sohbatpur district represents one of the most fragile surveillance environments in Balochistan, shaped by extreme poverty, flood vulnerability, and weak health infrastructure. Following the catastrophic floods of 2022, which displaced over 80% of the district's population, disease surveillance systems experienced prolonged disruption, with many Basic Health Units (BHUs) rendered nonfunctional (PDMA Balochistan, 2023; WHO, 2023). According to the National Institute of Health (NIH) Weekly Epidemiological Bulletin (Week 26, 2025), Sohbatpur reported persistently high incidences of Acute Watery Diarrhea, Malaria, and Suspected Measles, conditions closely linked to stagnant floodwaters, compromised sanitation, and limited access to safe drinking water (NIH, 2025b). Reporting compliance from public health facilities remained below provincial averages, reflecting shortages of trained surveillance staff and intermittent digital connectivity. Although emergency disease surveillance was temporarily strengthened through WHO-supported mobile health teams, the absence of a permanent laboratory facility and reliance on paper-based reporting continue to delay outbreak detection and response in the district (WHO Pakistan, 2024; Ullah et al., 2025).

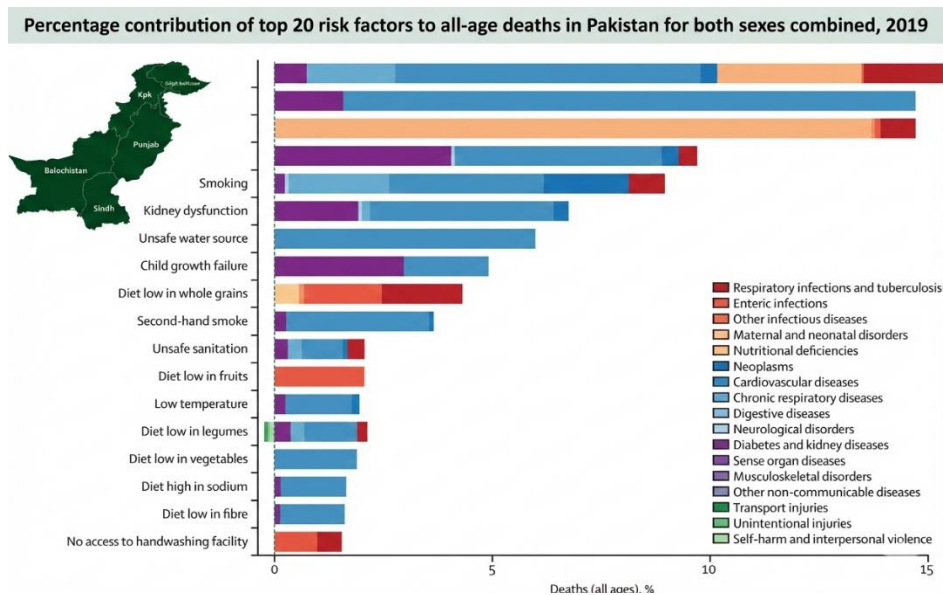


Figure 3: Contribution of Environmental and Behavioral Risk Factors to Mortality in Pakistan (2019)

### Case Studies: Ebola Virus Disease (EVD) in Africa

The response to Ebola outbreaks in Africa has served as a critical testing ground for modern surveillance and containment strategies. Gaps in early detection and management in West Africa during the 2013–2016 epidemic underscored the need for resilient, electronics-based reporting (Shannon, 2023).

#### The RITE Strategy in Liberia

In response to the West African outbreak, the CDC and partners implemented the Rapid Isolation and Treatment of Ebola (RITE) strategy in Liberia. This approach focused on rapid deployment of mobile teams to remote areas to isolate cases and trace contacts immediately. RITE successfully cut the time to control outbreaks in half and significantly increased survival rates by bringing care closer to the community (CDC, 2014).

#### Genomic Surveillance in the 10th DRC Outbreak

The 2018–2020 Ebola outbreak in the eastern DRC was the second-largest globally and occurred amidst chronic insecurity and community mistrust. Despite these challenges, decentralized field laboratories utilizing GeneXpert and NGS technology provided rapid, structured responses (ITG Research, 2024). Genomic data allowed for the detection of recrudescence where the virus remains hidden in immune-privileged niches like the testes and resurfaces months later, potentially sparking new transmission chains. This understanding shifted surveillance from purely reactive to long-term monitoring of survivors to prevent flare-ups (Mbala-Kingebeni et al., 2021).

#### Ethical Considerations, Data Privacy, and Governance

As surveillance systems become more data-intensive, the ethical and legal frameworks governing them must evolve to protect individual privacy while serving the public good (Syahmi, 2025)

#### Privacy-Preserving Technologies

The use of personal health data and digital exhaust raises significant privacy concerns. Federated Learning (FL) has emerged as a solution, allowing AI models to be trained on decentralized data from various hospitals without requiring the raw patient records to be shared or centralized (Karamat et al., 2025). Similarly, blockchain technology is being explored to ensure secure data

handling and transparency in early warning systems, preventing unauthorized access and maintaining a tamper-proof audit trail (Long et al., 2021).

### **Addressing Algorithmic Bias and Equity**

AI systems are only as good as the data they are trained on. There is a documented risk that algorithms trained on biased datasets may worsen social inequities. For example, a cost-based algorithm used in some healthcare systems was found to systematically underestimate the severity of illness in black patients, leading to reduced referrals for care (Abaoud et al., 2023). In LMICs, unrepresentative data and weak oversight can amplify these biases, making it essential to develop ethical and fair evaluation frameworks that include cultural adaptation and inclusive design (Woods et al., 2023).

### **Future Directions: 2026–2030**

The next five years will likely see the maturation of global intelligence networks and the deeper integration of AI into every layer of public health (George et al., 2024).

### **Global Pathogen Surveillance Networks**

The establishment of the WHO Pandemic Hub in Berlin aims to transform global surveillance into a collaborative intelligence ecosystem (Eurosurveillance, 2022). Key objectives include:

- Developing interoperable data platforms that preserve custodianship while promoting accessibility ((Vishal, 2025).
- Hosting the International Pathogen Surveillance Network (IPSN) to strengthen global genomic capacity ((Agbaakin et al., 2024).
- Utilizing human-driven, AI-enhanced insights to make complex epidemiological data accessible and actionable for decision-makers (WHO, 2024).

### **The Shift to Value and Outcomes**

By 2026, the digital health sector is expected to shift from demonstrating technical capability to delivering measurable clinical and financial ROI (S3 Connected Health, 2024). Remote patient monitoring (RPM) and ambient clinical intelligence will become standard components of healthcare, reduce physician cognitive load and allowing for proactive early intervention for both infectious and chronic conditions (Mia-Care, 2024).

### **2026 Public Health Data Strategy Milestones**

The CDC's upcoming goals for 2025–2026 emphasize expanding real-time access to emergency department (ED) visit data across at least 45 states and four territories. This strategy aims to grow the adoption of common legal agreements and standards like FHIR-based automated exchanges, ensuring that data moves faster than the pathogens it tracks (Dias, 2020). The continued evolution of modern epidemiological approaches necessitates a multidisciplinary effort, combining advanced analytics with community engagement and robust international governance (Bhutta et al., 2021). By leveraging the power of AI, genomics, and wastewater monitoring, the global public health community moves closer to a future where outbreaks are predicted and contained before they escalate into pandemics (Agbaakin et al., 2024). This paradigm shift requires not only technological innovation but also a renewed commitment to equity, ensuring that the most vulnerable populations in regions like Balochistan and the DRC benefit from the same high-tech early warning systems as those in more developed healthcare landscapes (Zinzuwadia et al., 2024). The integration of these tools into a One Health framework recognizing the interconnectedness of human, animal, and environmental health remains the ultimate goal for global health security (OHHLEP, 2022).

## Conclusions

The continued advancement of modern epidemiological approaches demands a synergistic integration of AI, genomics, WBE, and big data analytics with robust community engagement and international governance. By addressing infrastructure gaps in LMICs like Pakistan's Balochistan and learning from case studies such as Ebola in Africa, public health systems can achieve proactive outbreak containment. Ethical frameworks must evolve to mitigate biases and ensure privacy in data-intensive surveillance. Looking ahead to 2026–2030, the maturation of global networks like the WHO Pandemic Hub and a shift toward outcome-driven digital health will foster equitable access to early warning systems. Ultimately, embracing a One Health paradigm interconnecting human, animal, and environmental health represents the cornerstone for resilient global health security, enabling the prediction and mitigation of emerging threats before they escalate into pandemics.

## References

- Abaoud, M., Almuqrin, M. A., & Khan, M. F. (2023). Advancing federated learning through novel mechanism for privacy preservation in healthcare applications. *IEEE Access*, 11, 83562-83579.
- Abiuro, G. A., Mbera, G. B., & De Allegri, M. (2014). Gaps in universal health coverage in Malawi: a qualitative study in rural communities. *BMC health services research*, 14(1), 234.
- Agbaakin, O., & Iyorkar, V. (2024). Transforming global health through multimodal deep learning: Integrating NLP and predictive modelling for disease surveillance and prevention. *World Journal of Advanced Research and Reviews*, 24, 95-114.
- Aijaz, M., Ahmad, M., Ansari, M. A., & Ahmad, S. (2023). Antimicrobial resistance in a globalized world: current challenges and future perspectives. *Int. J. Pharm. Drug Des*, 1(1), 7-22.
- Ali, D. N. (2025). Planning and Implementing one Health Solutions in Pakistan. *Climate Change and Water-Related Challenges in Pakistan: Tangible Solutions*, 172.
- Amona, E. B., Sahoo, I., Su, Y., Boone, E. L., Nelis, G., & Ghanam, R. (2025). Essential Workers at Risk: An Agent-Based Model (SAFE-ABM) with Bayesian Uncertainty Quantification. arXiv preprint arXiv:2505.01243.
- Balochistan Post, The. (2017, November 20). *WHO report reveals 257 healthcare facilities in Balochistan nonexistent*. <https://thebalochistanpost.net/2024/04/who-report-reveals-257-healthcare-facilities-in-balochistan-nonexistent/>
- Barros, J. M., Duggan, J., & Rebholz-Schuhmann, D. (2020). The application of internet-based sources for public health surveillance (infoveillance): Systematic review. *Journal of medical internet research*, 22(3), e13680.
- Bashir, S., Khan, J., Danish, M., & Bashir, W. (2023). Governance and development challenges in Balochistan: A comparative study with other provinces and way forward. *International Journal of Contemporary Issues in Social Sciences*, 2(4), 620-649.
- Becker, J., Müller, B., & Grimm, V. (2024). The spatial integration of agent-based models and compartmental models for infectious disease modeling. *Journal of Artificial Societies and Social Simulation*, 28(1). <https://doi.org/10.18564/jasss.5394>
- Bhutta, M. N. M., & Ahmad, M. (2021). Secure identification, traceability and real-time tracking of agricultural food supply during transportation using internet of things. *Ieee Access*, 9, 65660-65675.
- Cazap, E., Magrath, I., Kingham, T. P., & Elzaway, A. (2016). Structural barriers to diagnosis and treatment of cancer in low-and middle-income countries: the urgent need for scaling up. *Journal of Clinical Oncology*, 34(1), 14-19.
- Centers for Disease Control and Prevention. (2014, October). *CDC response in heavily affected countries*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4622264/>

- Centers for Disease Control and Prevention. (2022, July). *Wastewater surveillance success stories*. <https://www.cdc.gov/advanced-molecular-detection/php/success-stories/wastewater-surveillance.html>
- Centers for Disease Control and Prevention. (2024). *Public health data strategy: PHDS milestones 2025 and 2026*. <https://www.cdc.gov/public-health-data-strategy/php/about/phds-milestones-2025-and-2026.html>
- Croce, D., Borazio, F., Gambosi, G., Basili, R., Margiotta, D., Scaiella, A., ... & Pilati, S. (2023). Intelligent Natural Language Processing for Epidemic Intelligence. *IJCoL. Italian Journal of Computational Linguistics*, 9(9-2).
- Cureus Journal of Medical Science. (2025). *Digital epidemiology and the use of non-traditional data sources*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12587752/>
- Dias, B. L. (2020). Big Data in Public Health: Real-Time Epidemiology Using Mobility and Environmental Data to Predict Outbreaks. *International Journal of Cell Science and Biotechnology*, 9(01), 05-10.
- Eurosurveillance Editorial Team. (2022). The WHO Hub for Pandemic and Epidemic Intelligence: Supporting better preparedness for future health emergencies. *Eurosurveillance*, 27(20).(<https://www.google.com/search?q=https://doi.org/10.2807/1560-7917.ES.2022.27.20.2205191>)
- Farkas, K., Williams, R. C., Hillary, L. S., Garcia-Delgado, A., Jameson, E., Kevill, J. L., ... & Jones, D. L. (2025). Harnessing the power of next-generation sequencing in wastewater-based epidemiology and global disease surveillance. *Food and Environmental Virology*, 17(1), 5.
- Gado, P., Oparah, O. S., Ezech, F. E., Gbaraba, S. V., Adeleke, A. S., & Omotayo, O. (2020). Framework for Developing Data-Driven Nutrition Interventions Targeting High-Risk Low-Income Communities Nationwide. *Framework*, 1(3).
- George, A. S., George, A. H., & Baskar, T. (2024). Artificial intelligence and the future of healthcare: Emerging jobs and skills in 2035. *Partners Universal Multidisciplinary Research Journal*, 1(1), 1-21.
- Giovanetti, M. (2023). Integrated genomic disease surveillance trends and challenges. *EASCI Journal*. <https://easci.eahealth.org/easci/article/view/106/568>
- Gordeev, V., Hölzer, M., Desirò, D., Goraichuk, I. V., Knyazev, S., Solo-Gabriele, H., ... & Mangul, S. (2025). Leveraging wastewater sequencing to strengthen global public health surveillance. *BMC global and public health*, 3(1), 1-5.
- Government of Balochistan. (2025, March). *Balochistan Health Institutions Reforms Act, 2025*.([https://pabalochistan.gov.pk/storage/8402/688c8b7c5dc7a\\_NOTIFICATION-OF-THE-BALOCHISTAN-HEALTH-INSTITUTIONS-REFORMS-ACT-2025-.pdf](https://pabalochistan.gov.pk/storage/8402/688c8b7c5dc7a_NOTIFICATION-OF-THE-BALOCHISTAN-HEALTH-INSTITUTIONS-REFORMS-ACT-2025-.pdf))
- Hooper Lundy & Bookman. (2024). *Predicting digital health trends in 2026*. <https://hooperlundy.com/predicting-digital-health-trends-in-2026/>
- ITG Research. (2024). *Implementation and evaluation of new tools for Ebola virus disease response during eastern Democratic Republic of Congo outbreaks*. <https://research.itg.be/en/publications/implementation-and-evaluation-of-new-tools-for-ebola-virus-diseas/>
- Karamat, F., Rahman, A. U., Saqia, B., Zafar, A., & Khan, W. A. (2025, April). Addressing Privacy-Preservation in Healthcare Using Federated Learning: A Survey. In *Artificial Intelligence and Applications*.
- Kim, Y., & Cho, N. (2022). A Simulation study on spread of disease and control measures in closed population using ABM. *Computation*, 10(1), 2.

- Lakmali, R. G. N., Genovese, P. V., & Abewardhana, A. A. B. D. P. (2024). Evaluating the efficacy of agent-based modeling in analyzing pedestrian dynamics within the built environment: A comprehensive systematic literature review. *Buildings*, 14(7), 1945.
- Lessani, M. N., Li, Z., Jing, F., Qiao, S., Zhang, J., Olatosi, B., & Li, X. (2024). Human mobility and the infectious disease transmission: a systematic review. *Geo-Spatial Information Science*, 27(6), 1824-1851.
- Lizot, L. F., Bastiani, M. F., de Souza, C. F., Hahn, R. Z., & Linden, R. (2023). Sampling techniques in wastewater-based epidemiology approach. In *Wastewater-Based Epidemiology for the Assessment of Human Exposure to Environmental Pollutants* (pp. 61-81). Academic Press.
- Long, G., Shen, T., Tan, Y., Gerrard, L., Clarke, A., & Jiang, J. (2021). Federated learning for privacy-preserving open innovation future on digital health. In *Humanity driven AI: productivity, well-being, sustainability and partnership* (pp. 113-133). Cham: Springer International Publishing.
- Mayaki, L. D. (2025). Artificial intelligence and predictive outbreak detection systems. *GSC Biological and Pharmaceutical Sciences*, 33(03), 001–014. <https://doi.org/10.30574/gscbps.2025.33.3.0484>
- Mbala-Kingebeni, P., Nsunda, B., & Gidoïn, C. (2021). *Progress in genomic surveillance over the course of the outbreak*. ResearchGate.([https://www.researchgate.net/figure/Progress-in-genomic-surveillance-over-the-course-of-the-outbreak-a-Total-numbers-of\\_fig1\\_350823563](https://www.researchgate.net/figure/Progress-in-genomic-surveillance-over-the-course-of-the-outbreak-a-Total-numbers-of_fig1_350823563))
- Mia-Care. (2024). *The top 5 upcoming trends in 2026 digital health*. <https://mia-care.io/industry-key-insights/top-5-upcoming-trends-2026-digital-health/>
- Mienye, I. D., & Swart, T. G. (2024). A comprehensive review of deep learning: Architectures, recent advances, and applications. *Information*, 15(12), 755.
- National Institute of Health, Islamabad. (2024c, December). *Public health bulletin - Pakistan, week 51, 2024*.([https://www.nih.org.pk/wp-content/uploads/2025/01/Weekly\\_Report-51-2024.pdf](https://www.nih.org.pk/wp-content/uploads/2025/01/Weekly_Report-51-2024.pdf))
- National Institute of Health, Islamabad. (2025b, June). *Public health bulletin - Pakistan, week 26, 2025*.([https://www.nih.org.pk/wp-content/uploads/2025/08/Weekly\\_Report\\_26-2025.pdf](https://www.nih.org.pk/wp-content/uploads/2025/08/Weekly_Report_26-2025.pdf))
- One Health High-Level Expert Panel. (2022). *OHHLEP report 2022*. Food and Agriculture Organization. <https://cdn.who.int/media/docs/default-source/one-health/ohhlep/ohhlep-report-2022.pdf>
- Pakistan Journal of Medical Research Editorial Board. (2024, January). Integrated disease surveillance and response system (IDSRS) in Pakistan: Challenges and way forward. *Pakistan Journal of Medical Research*, 63(1). <https://pjmr.org.pk/index.php/pjmr/article/view/740>
- Panagiotou, N., Katakis, I., & Gunopulos, D. (2016). Detecting events in online social networks: Definitions, trends and challenges. *Solving Large Scale Learning Tasks. Challenges and Algorithms: Essays Dedicated to Katharina Morik on the Occasion of Her 60th Birthday*, 42-84.
- PMC11717196. (2024). *Metapopulation graph transformer neural network (M-Graphormer) for high-dimensional parameter estimation and multi-regional epidemic prediction*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11717196/>
- PMC12230060. (2025). *Privacy-preserving technologies in infectious disease surveillance*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12230060/>
- Psomopoulos, F., O’Cathail, C., Anastasiadou, N., Cocuzza, C., Cuypers, W. L., Idowu, O., ... & ELIXIR Wastewater Surveillance Working Group. (2025). Toward a unified approach:

- Considerations for bioinformatic and sequencing activities & data in wastewater surveillance of biologic public health threats. *Open Research Europe*, 5, 267.
- Raisani, N. M. J. K., & Singhaputargun, N. (2025). Understanding Conflict and Reconciliation: Lessons from Balochistan, Pakistan. *Connexion: Journal of Humanities and Social Sciences*, 14(1), 1-26.
- Rony, M. M. A., Soumik, M. S., & Akter, F. (2023). Applying Artificial Intelligence to Improve Early Detection and Containment of Infectious Disease Outbreaks, Supporting National Public Health Preparedness. *Journal of Medical and Health Studies*, 4(3), 82-93.
- S3 Connected Health. (2024). *The key trends that will define digital health in 2026*. <https://www.s3connectedhealth.com/blog/the-key-trends-that-will-define-digital-health-in-2026>
- Salam, M. A., Al-Amin, M. Y., Salam, M. T., Pawar, J. S., Akhter, N., Rabaan, A. A., & Alqumber, M. A. (2023, July). Antimicrobial resistance: a growing serious threat for global public health. In *Healthcare* (Vol. 11, No. 13, p. 1946). MDPI.
- Shannon II, F. Q. (2023). Public health evaluation of the Ebola virus disease surveillance and response system in Liberia (Doctoral dissertation, University of Otago).
- Syahmi, H. (2025). Privacy and ethical implications of big data utilization in public transportation surveillance. *International Journal of Advanced Cybersecurity Systems, Technologies, and Applications*, 9(1), 1-10.
- Tzavellas, N. P., Atzemoglou, N., Bozidis, P., & Gartzonika, K. (2026). The Role of Environmental and Climatic Factors in Accelerating Antibiotic Resistance in the Mediterranean Region. *Acta Microbiologica Hellenica*, 71(1), 1.
- Ullah, S., Ullah, N., Tariq, A., Panezai, S., Tahir, P., & Khan, B. (2025). Spatial assessment of the availability of healthcare facilities at district level in Balochistan using GIS: identifying gaps and way forward. *Spatial Information Research*, 33(2), 8.
- Villanueva-Miranda, I., Xiao, G., & Xie, Y. (2025). Artificial Intelligence in Early Warning Systems for Infectious Disease Surveillance: A Systematic Review. *Frontiers in Public Health*, 13, 1609615.
- Vishal, K. (2025). Natural Language Processing in Digital Health: Transforming Clinical Narratives into Actionable Intelligence.
- Woods, W. A., Watson, M., Ranaweera, S., Tajuria, G., & Sumathipala, A. (2023). Underrepresentation of low and middle income countries (LMIC) in the research literature: ethical issues arising from a survey of five leading medical journals: have the trends changed?. *Global public health*, 18(1), 2229890.
- World Health Organization. (2024). *WHO Hub for Pandemic and Epidemic Intelligence*. <https://pandemichub.who.int/>
- Wu, N., Green, B., Ben, X., & O'Banion, S. (2020). Deep transformer models for time series forecasting: The influenza prevalence case. arXiv preprint arXiv:2001.08317.
- Zhang, W., Valencia, A., & Chang, N. B. (2021). Synergistic integration between machine learning and agent-based modeling: A multidisciplinary review. *IEEE Transactions on Neural Networks and Learning Systems*, 34(5), 2170-2190.
- Zhu, W., Wang, D., Li, P., Deng, H., & Deng, Z. (2025). Advances in Wastewater-Based Epidemiology for Pandemic Surveillance: Methodological Frameworks and Future Perspectives. *Microorganisms*, 13(5), 1169.
- Zinzuwadia, A. N., Mineeva, O., Li, C., Farukhi, Z., Giulianini, F., Cade, B., ... & Demler, O. (2024). Tailoring risk prediction models to local populations. *JAMA cardiology*, 9(11), 1018-1028.