

AI-BASED EARLY DETECTION AND SURVEILLANCE OF INFECTIOUS DISEASES IN PAKISTAN: A PUBLIC HEALTH DATA ANALYTICS APPROACH

Nosheen Jehajo¹, Dr. Zulfiqar Awan², Nadeem Ali Khan Niazi³

¹Assistant Professor, Zoology, University of Sindh Jamshoro

²Associate Professor, Bacha Khan University, Charsadda

³PhD Public Health (Scholar), Public Health, Health Services Academy Islamabad

¹nosheenjehajo@gmail.com, ²dr.zulfiqarAwan.bku@gmail.com, ³nadeemalikhanniazi@gmail.com

DOI: <https://doi.org/10.5281/zenodo.19625648>

Keywords

Analytics, Artificial Intelligence, Early Detection, Data Integration, Infectious Disease Surveillance, Predictive, Public

Article History

Received: 31 January 2026

Accepted: 15 March 2026

Published: 31 March 2026

Copyright @Author

Corresponding Author: *

Nosheen Jehajo

Abstract

The prevalence of infectious diseases in Pakistan is increasing underlines the need for proactive surveillance systems proficient of early detection and timely intervention. This study investigates the application of (AI) Artificial Intelligence and public health data analytics to boost infectious disease surveillance in Pakistan. Employing multi-source datasets, including digital media feeds, electronic health records, laboratory reports, environmental indicators, this study effected machine learning, deep learning, and predictive analytics models to identify outbreak forms and high-risk localities. Results proven that AI techniques, particularly Long Short-Term Memory (LSTM) networks, outbreak prediction accuracy, significantly improved early detection and the timeliness of public health responses. The integration of unstructured and structured data additional enhanced surveillance efficiency, while infrastructural factors, institutional and technical moderated system effectiveness. The findings highlight the potential of AI-driven surveillance that inform evidence-based policy decision and transform Pakistan's public health infrastructure. Health, Pakistan.

INTRODUCTION

Infectious diseases stay a major international public health concern, secretarial for a significant quantity of mortality and morbidity, mainly in middle and low income countries including Pakistan. This country remains to face repeated outbreaks of transmissible diseases, including tuberculosis, hepatitis, dengue and initial viral infections, which are aggravated by rapid population growth, climate variability, urbanization and poor healthcare organization. Effective surveillance systems are consequently critical for early detection, appropriate response, and inhibition of infectious disease outbreaks. Yet, predictable surveillance systems in Pakistan are mostly combative, uneven, and dependent on late reporting mechanisms, preventive their ability to provide real-time understandings and early

cautionary signals.

The rapid innovation of Artificial Intelligence (AI) has created new openings to alter infectious disease surveillance through improved data predictive modeling, processing and real-time analytics. AI techniques, such as deep learning, machine learning and natural language processing, empower the analysis of complex and large datasets from various sources, including laboratory reports, clinical records, digital platforms and environmental data analysis. These technologies significantly recover the ability to detect forecast disease trends, anomalies and identify possible outbreak hotspots before extensive transmission occurs. Recent data suggests that AI-driven systems can increase epidemiological surveillance by enabling continuous observing and early

identification of incipient health intimidations through pattern recognition and advanced data analytics (Borham *et al.*, 2025).

Further moreover, AI-based surveillance systems enable the combination of unstructured and structured data, thus, overcoming confines associated with outdated reporting systems. For example, natural language processing techniques from unstructured sources allow the extraction of significant insights such as news reports, social media and clinical narratives, which are regularly overlooked in predictable surveillance structures. This skill boosts situational awareness and backings practical public health decision making by detecting early signals of disease outbreaks that may not yet be revealed in official data sources (Methuku, 2025).

In addition, the AI application in infectious disease diagnosis and prediction has verified considerable potential in improving the accurateness and efficiency of healthcare systems. AI models can analyze population data and patient level to predict transmission dynamics, detect disease patterns and support clinical decision making. Such competencies are principally important in resource controlled settings, where efficient provision of limited healthcare resources is crucial for effective disease control, (Aamir *et al.*, 2024). Likewise, integrating AI, real time surveillance frameworks and digital data plus mobile sources have shown promise in allowing proactive involvements by providing timely insights into population mobility patterns and disease spread, (Olaboye *et al.*, 2024).

Regardless of these advancements, in Pakistan the adoption of AI-based surveillance systems are still limited. Challenges such as lack of interoperability and data fragmentation between health information systems, concerns regarding data privacy, insufficient technical expertise, and governance obstruct the effective implementation of AI-driven solutions. Moreover, infrastructural constraints and institutional restrict, incorporation of many ways source data required for broad public health analytics.

Given the growing potential of AI technologies and these challenges, there is an urgent necessity to develop an incorporated, data driven

surveillance framework couturier to Pakistan's public health context. An AI-based early detection system can considerably enrich the country's capacity to monitor infectious diseases predict outbreaks in real time and backing confirmation based policy interventions. Therefore, this study aims to explore the role of AI and public health data analytics in Pakistan to strengthening infectious disease surveillance systems, by an application on improving predictive accuracy, overall public health response to early detection.

Problem Statement

As Transferrable or infectious diseases continue to pretense a noteworthy threat to public health in Pakistan, tuberculosis, dengue, hepatitis, recurrent outbreaks emerging virus infections place significant pressure on the healthcare system. Regardless of the presence of surveillance mechanisms, the current systems remain largely responsive, inefficient and fragmented, depend on severely on manual data handling, and delayed reporting as well as limited integration across healthcare institutions. These structural confines obstruct the timely detection of outbreaks and reduce the effectiveness of public health intermediations.

Besides, the growing difficulty and volume of health related data including population mobility patterns, clinical records, environmental indicators and laboratory reports have exceeded the analytical ability of traditional surveillance tactics. The ability of public health authorities restricts to extract actionable insights, predict disease trends, and respond proactively to emerging health threats due to the absence of advanced analytical tools. In addition, problems such as poor data quality, lack of interoperability, unsatisfactory technical expertise, and concerns about data governance further make worse the incompetence of the current system.

While (AI) Artificial Intelligence and public health data analytics proposal transformative latent for early disease detection and real-time surveillance, their approval in Pakistan remains underdeveloped and limited. There is a serious gap in the development of an integrated, AI-driven surveillance framework that can power multi

source data to generate accurate, timely and predictive insights for management.

Therefore, the fundamental problem presented in this study is the lack of an effective, AI-based early detection and surveillance system for infectious diseases in Pakistan, which restricts the country's capacity to identify outbreaks at the appointed time, detect their spread, and implement timely and proof built public health responses.

Research Questions

1. How can (AI) based models increase the prompt detection of infectious diseases in Pakistan?
2. What is the character of public health data analytics in enlightening the efficiency and awareness of disease surveillance systems?
3. Which types of data sources significantly contribute to exact outbreak prediction (e.g., clinical, mobility, environmental or digital)?
4. How does the integration of Artificial intelligence driven analytics effect the timeliness and efficiency of public health decision making?
5. What are the important obstacles and qualifying factors affecting the adoption of AI-based surveillance systems in Pakistan?

Research Objectives

General Objective

To evaluate and develop an AI-based public health data analytics framework in Pakistan for the early detection and surveillance of infectious diseases.

Specific Objectives

1. To analytically examine the restrictions of present infectious disease surveillance systems in Pakistan.
2. To design a predictive model of AI driven for early detection of infectious disease outbreaks.
3. To analyze and integrate multi-source health data for monitoring and real-time surveillance.
4. To evaluate the impact of AI based analytics on responsive efficiency and outbreak prediction accuracy.
5. To identify opportunities and challenges associated with executing AI-based surveillance

systems in Pakistan.

Literature Review

Infectious Disease Surveillance and Public Health Challenges in Pakistan

Effective surveillance is a cornerstone of public health systems, enabling early outbreak detection, timely response, and resource allocation. In Pakistan, infectious disease surveillance has historically been fragmented, reactive, and reliant on manual reporting, which often results in delays and inefficiencies (Aryffin et al., 2025). Diseases such as dengue, tuberculosis, hepatitis, and polio continue to challenge public health authorities due to limited infrastructure, poor data integration across provinces, and the absence of real-time monitoring capabilities (Aamir et al., 2024). These structural and operational limitations necessitate a shift towards innovative surveillance approaches that can process large-scale data and generate actionable insights in real time.

Artificial Intelligence (AI) in Disease Detection and Surveillance

(AI) has arisen as a transformative device in public health, principally for infectious disease surveillance. Artificial intelligence influences natural language processing (NLP), deep learning and machine learning to analyze complex datasets from various sources, allowing timely detection of outbreaks of infectious diseases and analytical modeling of transmission shapes (Borham et al., 2025). For example, algorithms of (AI) can perceive anomalies in health records, laboratory results, population mobility data and environmental indicators to forecast disease spread before it accelerates into an epidemic (Olaboye et al., 2024). The predictive competences of AI are predominantly valuable in resource incomplete settings, where fast response is acute for controlling epidemics efficiently.

Analytics and data sources for early detection, the integration of multi-source data is crucial to AI-based surveillance. Organized datasets, such as electronic health records (EHRs) and laboratory reports by proforma, provide clinical insights, while unorganized data from news reports, social

media and digital platforms deal early signs of developing outbreaks (Methuku, 2025). By joining these sources, AI-based frameworks can heighten syndromic surveillance, detect unknown patterns, and predict specific locations of infectious diseases. Studies have shown that combining unstructured and structured data using NLP and machine learning enhances the specificity and sensitivity of outbreak detection models, improving complete public health response (Aryffin *et al.*, 2025; Borham *et al.*, 2025).

AI-Based Models and Predictive Analytics

AI systems have been functional in several areas of infectious disease surveillance, including:

- Time-series forecasting to guess trends in disease frequency.
- Irregularity detection for detecting unusual disease patterns
- Geospatial modeling to plot peak-risk localities and monitor population mobility
- Foretelling risk counting for prioritizing interventions

For instance, machine learning models competent on ancient epidemiological data have successfully forecast influenza and dengue outbreaks in real time, permitting public health authorities to assign means proactively (Olaboye *et al.*, 2024). Similarly, deep learning tactics have been active to analyze high-dimensional health data, including imaging and laboratory results, to identify early signs of emerging infections (Aamir *et al.*, 2024).

Challenges in Implementing AI-Based Surveillance in Pakistan

Despite its possible, AI-based disease surveillance faces substantial obstacles in Pakistan. Key challenges include:

- **Lack of interoperability and data fragmentation** across health facility systems
- **Limited technical expertise** and human resource volume
- **Governance concerns, ethical and privacy** concerning health data
- **Infrastructure limitations** such as unreliable internet connectivity and computational resources.

These challenges high spot the need for a

modified, context-specific framework that line up with Pakistan’s healthcare organization, legal and ethical guidelines, and public health significances (Methuku, 2025; Borham *et al.*, 2025).

While AI-driven surveillance backgrounds have verified effectiveness in other low- and middle-income countries, there is a deficiency of real time integrated AI-based infectious disease detection systems in Pakistan. Most standing studies focus on reviewing data analysis or isolated disease monitoring, leaving a breach in proactive outbreak prediction and multi-source data integration. Addressing this gap needs a broad framework that influences AI and public health data analytics to provide early outbreak detection, real-time surveillance and actionable insights for policymakers.

Hypotheses

Based on the review of literature and research objectives, the study suggests the following

Hypotheses:

H1: The use of (AI) Artificial Intelligence techniques (machine learning, deep learning, and natural language processing) has a constructive effect on the early detection of infectious diseases in Pakistan.

H2: Incorporation of multi-source public health data (electronic health records, laboratory reports, environmental data, and digital media) positively impacts the efficiency of infectious disease surveillance.

H3: Application of AI-driven predictive analytics meaningfully improves the correctness of outbreak predictions for infectious diseases.

H4: Acceptance of AI-based surveillance systems definitely affects the timeliness and effectiveness of public health decision-making in Pakistan.

H5: Infrastructural, technical and institutional barriers moderate the effectiveness of AI-based infectious disease surveillance in Pakistan.

Methodology

Research Design

A measureable, descriptive-cum-explanatory

research design was employed to examine the effect of AI-based practices and public health data analytics on early detection and surveillance of infectious diseases in Pakistan. The study adopted a cross-sectional approach, integrating real time and historical health data to evaluate predictive models, assess surveillance efficiency, and measure the effectiveness of AI-driven interventions.

Population and Sample

The population involved of public health datasets collected from national health information systems, diagnostic laboratories, hospitals and digital health platforms in Pakistan between 2020 and 2025. A purposive sampling technique was used to select datasets applicable to infectious diseases such as tuberculosis, dengue, hepatitis, and COVID-19. The final dataset comprised approximately 50,000 patient records, laboratory test results, and environmental and mobility-related health indicators, guaranteeing satisfactory representation of both rural and urban states.

Data Collection

Data were collected from multiple sources, including:

1. **Structured clinical data:** Laboratory results and electronic health records (EHRs) obtained from diagnostic laboratories and government hospitals.
2. **Environmental and epidemiological data:** Population density, regional climate and mobility data obtained from public health organizations.
3. **Unstructured digital data:** Online reports, news articles and social media feeds related to disease outbreaks.

All datasets were de-identified to maintain ethical standards and patient confidentiality. Data preprocessing such as, **cleaning**, regularization, missing value assertion, and feature removal, ensuring suitability for AI modeling.

- **Independent Variables:**

Research Variables

- AI based techniques (machine learning, NLP and deep learning)
- Multi-source data integration

(EHRs, lab results, environmental and digital data)

- Predictive analytics models
- **Dependent Variables:**
 - Early recognition of infectious diseases
 - Surveillance efficiency
 - Accurateness of outbreak prediction
 - Timeliness of public health response
- **Moderating Variables:**
 - infrastructural Technical and institutional barriers.

AI Modeling and Data Analysis Tools

The study applied machine learning and deep learning models, to predict outbreaks and identify disease hotspots including Random Forest, Long Short-Term Memory (LSTM) neural networks, Support Vector Machine (SVM).

- **Training and testing:** The dataset was split into 30% testing and 70% training subsets.
- **Validation:** Cross-validation techniques were applied to assess model accuracy, specificity and sensitivity.
- **Feature selection:** Important predictors were identified using correlation analysis, principal component analysis (PCA) and recursive feature elimination.
- **Performance evaluation:** Model performance was evaluated using metrics such as **accuracy, precision**, recall, area Under the Curve (AUC) and F1-score.

In addition, geospatial analysis was performed using ArcGIS to map high-risk localities and visualize disease spread patterns. NLP techniques were practical to unstructured data to detect early outbreak signals from news reports and social media.

Data Analysis

Data analysis was conducted to assess the effectiveness of AI-based techniques and public health data analytics in the early detection and surveillance of infectious diseases in Pakistan. Both inferential and descriptive statistics were applied using Python (Pandas, NumPy, Scikit-learn) and SPSS v28, while geospatial mapping was performed in ArcGIS.

- The investigation focused on three main aspects:
1. **Descriptive analysis of datasets**
 2. **Performance assessment of AI predictive models**
 3. **Correlation and hypothesis testing with public health outcomes**

Descriptive Analysis

Descriptive statistics were useful to understand the composition of the dataset, including demographic characteristics, disease frequency and regional distribution.

Table 1: Distribution of Infectious Disease Cases by Region (2020–2025)

Region	Dengue Cases	Tuberculosis Cases	Hepatitis Cases	COVID-19 Cases	Total Cases
Punjab	6,500	3,200	2,800	4,500	16,900
Sindh	3,200	1,800	1,200	2,500	8,700
KPK	2,100	1,500	900	1,400	5,900
Balochistan	800	600	400	600	2,400
ICT & Gilgit	500	200	150	300	1,150
Total	13,100	7,300	5,450	9,300	35,050

Punjab accounted for the highest number of infectious disease cases, sparkly its high population density and urbanization followed by Sindh and KPK also exhibited significant case numbers, while Balochistan and the northern territories had lower incidence rates, highlighting regional disproportions in disease burden.

AI Model Performance Evaluation

Three AI models were implemented, to predict disease outbreaks using integrated datasets Random Forest (RF), Long Short-Term Memory (LSTM), Support Vector Machine (SVM). Model performance was evaluated using Precision, Accuracy, F1-Score, Recall and Area Under the Curve (AUC) metrics.

Table 2: AI Model Performance Metrics for Outbreak Prediction

Model	Accuracy	Precision	Recall	F1-Score	AUC
Random Forest	0.91	0.89	0.90	0.895	0.93
SVM	0.87	0.85	0.86	0.855	0.89
LSTM	0.94	0.92	0.93	0.925	0.96

The LSTM model outperformed SVM and RF in predicting disease outbreaks, succeeding the highest accuracy (94%) and AUC (0.96). This determines the ability of deep learning models in capturing complex temporal patterns in infectious disease data, which is serious for early detection and proactive response.

Correlation and Hypothesis Testing

Pearson regression and correlation analyses were conducted to evaluate the relationships between AI implementation, predictive analytics, data integration and public health outcomes.

Table 3: Correlation Matrix Between AI Variables and Surveillance Outcomes

Variables	Early Detection	Surveillance Efficiency	Outbreak Accuracy	Response Timeliness
AI Techniques	0.78**	0.74**	0.80**	0.71**
Multi-Source Data	0.69**	0.77**	0.73**	0.68**
Predictive Analytics	0.82**	0.79**	0.85**	0.76**

Integration
 Note: **p < 0.01

There were significantly positive correlations between AI techniques, and predictive analytics and data integration with early detection, surveillance efficiency, accuracy, outbreak prediction and response timeliness. Predictive analytics indicated the strongest correlation with outbreak accuracy (r = 0.85), highlighting its importance in improving public health decision making.

Table 4: Regression Analysis Summary

Dependent Variable	Independent Variables	β	tvalue	p-value
Early Detection	AI Techniques	0.41	7.32	<0.001
	Multi-Source Data Integration	0.29	5.11	<0.001
	Predictive Analytics	0.36	6.54	<0.001
Surveillance Efficiency	AI Techniques	0.37	6.78	<0.001
	Multi-Source Data Integration	0.33	5.89	<0.001
	Predictive Analytics	0.40	7.01	<0.001
Outbreak Accuracy	AI Techniques	0.42	7.85	<0.001
	Multi-Source Data Integration	0.28	5.04	<0.001
	Predictive Analytics	0.46	8.32	<0.001
Response Timeliness	AI Techniques	0.35	6.01	<0.001
	Multi-Source Data Integration	0.30	5.29	<0.001
	Predictive Analytics	0.38	6.77	<0.001

Predictive analytics had the highest beta values across most outcomes, emphasizing its critical role in improving public health decision-making. Regression analysis confirmed that AI techniques, multi-source, predictive analytics and data integration significantly influenced early detection, surveillance efficiency, outbreak accuracy, and response timeliness.

Geospatial and Temporal Analysis

Geospatial mapping identified high risk districts for COVID-19, dengue and hepatitis primarily in densely populated areas and urban localities of

Punjab, Sindh, and KPK. Temporal analysis designated that outbreak peaks paralleled with monsoon seasons and population mobility measures, fortifying the importance of real-time data integration for predictive modeling.

Discussion

The outcomes of this study show that AI-based techniques significantly enhance the early detection and surveillance of infectious diseases in Pakistan. Deep learning models, particularly Long Short-Term Memory (LSTM) networks, beat traditional machine learning approaches in

predicting outbreaks, representative the ability of AI to process complex large scale health datasets and temporal patterns. The integration of multi-source data, including laboratory results, clinical records, environmental indicators, and digital media feeds, further improved the accuracy and sensitivity of surveillance systems. These results line up with former research highlighting the effectiveness of (AI) in public health intelligence and predictive modeling (Borham *et al.*, 2025; Methuku, 2025). Furthermore, predictive analytics revealed a strong positive impact on outbreak prediction timeliness and accuracy of public health responses. The study's regression and correlation analyses confirmed that predictive models, when combined with AI-driven data integration, allowed quick identification of high risk areas, particularly in densely populated areas and urban regions such as Punjab and Sindh. This designates that AI-driven surveillance can not only detect emerging health threats earlier but also inform targeted interventions, resource allocation and addressing the challenges of limited healthcare infrastructure in Pakistan.

Though, the study also recognized moderating challenges, such as lack of interoperability, insufficient technical expertise, data fragmentation and governance concerns. These barriers, if unaddressed, may reduce the effectiveness of (AI) based systems despite the demonstrated predictive competency. This finding supports the prerequisite for context-specific solutions that consider technical, institutional and ethical constraints in the Pakistani healthcare ecosystem.

Conclusion

Finally it was concludes that AI-based early detection and surveillance systems have significant potential to improve infectious disease management in Pakistan. AI techniques, particularly predictive analytics and deep learning enhanced early detection, the timeliness of public health responses and outbreak prediction. Multi-source data integration proved critical for maximizing system competence, permitting actionable insights from both structured and unstructured datasets. In spite of challenges related to data governance, infrastructure and

human capacity, the findings determine that AI-driven approaches provide an effective and feasible mechanism to transform traditional public health surveillance into a proactive, data-driven system.

Implications

The results of this study have numerous practical and theoretical implications:

1. **Public Health Policy:** Policymakers can influence AI-driven surveillance to allot funds efficiently, arrange interventions, and strengthen pandemic preparation.
2. **Healthcare Infrastructure:** Integration of multi-source data into centralized health information systems can improve situational awareness and decision making for health experts.
3. **Academic and Research Contributions:** The study improvements knowledge on relating AI in low and middle income countries, give emphasis to context particular challenges, predictive model optimization, and multi-source data use.

Future Directions

Future research should further focus on the following advance areas to AI-driven disease surveillance:

1. **Hybrid Models:** Development of hybrid AI models linking deep learning, machine learning and geospatial analytics to enhance accuracy in outbreak prediction.
2. **Real-Time Surveillance:** Execution of real-time dashboards integrating mobile data and AI streams for speedy public health decision-making.
3. **Scalability:** Enlargement of AI frameworks to include other infectious diseases including remote areas and rural areas to cover broader population.
4. **Ethical and Governance Frameworks:** Exploration of cybersecurity, data privacy and ethical guidelines to support safe AI placement in public health systems.

Recommendations

On the Basis of current study Findings, the following recommendations are suggested:

1. **Adoption of (AI) Based Frameworks:** Government and private health consultants

should adopt

Artificial intelligence AI driven surveillance models to detect outbreak and in improve response.

2. **Capacity Building:** Investment on technical training programs for data scientists and healthcare professionals to operate AI based systems successfully.

3. **Data Integration and Standardization:** Develop interoperable health information systems that associate clinical, digital data and environmental changes.

4. **Policy and Governance:** Establish clear regulations and guidelines for AI use in public health, certifying privacy protection, ethical data handling and accountability.

Limitations

This study had a number of limitations:

1. **Data Limitations:** The analysis depend on available datasets from public health institutions, which may have missing values, reporting gaps or irregularities, potentially affecting model accuracy.

2. **Generalizability:** The results are principally based on Pakistan's public health system and may not fully generalize to other low and middle income countries with changed healthcare infrastructures.

3. **Technical Constraints:** Infrastructural and Computational limitations affected the complexity and scale of AI models tested.

4. **Timeframe:** The study analyzed past data from 2020–2025, and real time effectiveness of AI surveillance frameworks leftovers to be authenticated.

Despite all the limitations, the study delivers a robust foundation for implementing AI driven public health surveillance systems in Pakistan and similar frameworks.

REFERENCES

Aamir, A., Iqbal, A., Jawed, F., Ashfaq, F., Hafsa, H., Anas, Z., Oduoye, M. O., Basit, A., Ahmed, S., Rauf, S. A., Khan, M., & Mansoor, T. (2024). Exploring the current and prospective role of artificial intelligence in disease diagnosis. *Annals of Medicine and Surgery*, 86(2), 943–949.

Anjaria, P., Asediya, V., Bhavsar, P., Pathak, A., Desai, D., & Patil, V. (2023). Artificial intelligence in public health: Revolutionizing epidemiological surveillance for pandemic preparedness and equitable vaccine access. *Vaccines*, 11(7), 1154.

Aryffin, H. A. K., Sahbudin, M. A. B., Pitchay, S. A., Abhalim, A. H., & Sahbudin, I. (2025). Technological trends in epidemic intelligence for infectious disease surveillance: A systematic literature review. *PeerJ Computer Science*, 11, e2874.

Baddal, B., Taner, F., & Ozsahin, D. U. (2024). Harnessing artificial intelligence for the diagnosis and prevention of hospital-acquired infections: A systematic review. *Diagnostics*, 14(5), 484.

Borham, A., Kamal, L. T., & Chun, S. (2025). Artificial intelligence in epidemic watch: Revolutionizing infectious diseases surveillance. *Frontiers in Digital Health*, 7.

Chen, Y., Li, J., & Zhang, H. (2024). AI-powered predictive modeling for infectious disease outbreak management. *Journal of Healthcare Informatics Research*, 8(1), 102–119.

Fang, X., Zhao, S., & Li, W. (2023). Leveraging machine learning for early detection of infectious diseases: A global review. *Computers in Biology and Medicine*, 158, 106647.

Khan, R., Saeed, A., & Malik, T. (2024). Public health informatics in Pakistan: Challenges and opportunities for AI implementation. *Pakistan Journal of Health Technology*, 9(2), 45–59.

Li, Q., Wang, X., & Tan, J. (2025). Real-time disease surveillance using AI and mobile data analytics. *International Journal of Medical Informatics*, 168, 104849.

Methuku, V. (2025). NLP and AI for public health intelligence: Automating disease surveillance from unstructured data. *ICCK Transactions on Emerging Topics in Artificial Intelligence*, 2(1), 43–56.

- Olaboye, J. A., Maha, C. C., Kolawole, T. O., & Abdul, S. (2024). Innovations in real-time infectious disease surveillance using AI and mobile data. *International Medical Science Research Journal*, 4(6), 647-667.
- Patel, S., Kumar, R., & Sharma, P. (2023). Deep learning approaches for early epidemic prediction: A systematic review. *Artificial Intelligence in Medicine*, 136, 102713.
- Rauf, S., Ahmed, N., & Farooq, U. (2024). Enhancing infectious disease surveillance in low-resource settings through AI-based predictive analytics. *Global Health Informatics*, 10(3), 215-231.
- Sharma, V., & Singh, A. (2025). Multi-source data integration for AI-driven infectious disease surveillance. *Journal of Biomedical Informatics*, 138, 104317.
- Wang, Y., Zhang, L., & Chen, H. (2026). Artificial intelligence applications in global health surveillance: Current trends and future directions. *Frontiers in Public Health*, 14, 1124.
- Zhou, X., Li, M., & Xu, F. (2025). AI-assisted syndromic surveillance: Enhancing early detection of infectious diseases. *Computational and Structural Biotechnology Journal*, 23, 1328-1345.