



Effectiveness of Artificial Intelligence-assisted Tools in Outbreak Detection, Surveillance and Emergency Response: A Systematic Review and Meta-analysis

Radhay Shyam^{1*}, Anna Nathan-Dagah¹, Anastesia Adaolisa Ekwe¹, Joseph Tikai¹, Sajos Yusuf Iliya¹, Edoki Samuel Siyesam¹, Selamawit Gebreyesus¹ and R. Aggarwal² 

¹Center for Sustainable and Health Research (CSHR) Foundation, Nayagaon, Near PGIMER, Chandigarh-160103, Punjab, India; cshrfoundation@outlook.com

²Department of Community and Family Medicine, Mandi Dabwali Road, Bathinda-151001, Punjab, India

Abstract

Background: The increasing frequency of emerging infectious diseases has exposed critical limitations in traditional surveillance systems. Artificial Intelligence (AI)-assisted tools have been proposed as scalable solutions to enhance outbreak detection and response. **Objective:** This systematic review and meta-analysis aimed to evaluate the performance of AI-assisted tools in infectious disease outbreak detection, surveillance, and emergency response. **Methods:** A systematic review was conducted following PRISMA guidelines. PubMed, Scopus, Web of Science, and IEEE Xplore were searched for studies published between 2011 and 2026. Quantitative studies reporting diagnostic or predictive performance metrics were eligible. Pooled accuracy, sensitivity, and area under the receiver operating characteristic curve (AUC) were summarised descriptively. **Result:** A total of included studies demonstrated high overall performance of AI/ML models in infectious disease surveillance. The pooled accuracy was 91.5% (95% CI: 88%–94%), with sensitivity of 0.82 (95% CI: 0.78–0.86), specificity of 0.92 (95% CI: 0.88–0.96), and AUC of 0.93 (95% CI: 0.90–0.95). While sensitivity showed moderate variability, accuracy, specificity, and AUC remained consistently high across studies. Moderate heterogeneity was observed, largely attributable to differences in study design, data sources, and geographic settings. Models using multimodal and image-based data achieved the highest performance, whereas studies in resource-limited settings reported comparatively lower sensitivity. Overall, AI/ML approaches demonstrated strong discriminative ability and improved timeliness, supporting their applicability for real-world infectious disease surveillance. **Conclusion:** AI/ML models demonstrate high accuracy, specificity, and overall discriminative performance in infectious disease surveillance, with added benefits in timeliness. Despite moderate heterogeneity across settings and methodologies, the findings support their robustness and potential for real-world implementation. Future efforts should focus on improving sensitivity in resource-limited settings, standardising methodologies, and enhancing model generalisability across diverse populations and data sources.

Keywords: Artificial Intelligence, Infectious Disease Surveillance, Machine Learning, Meta-Analysis, Systematic Review

Article Received on: 13-07-2026

Revised on: 14-08-2026

Accepted on: 15-09-2026



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0.

*Author for correspondence

1. Introduction

1.1 Background and Rationale

Surveillance is the backbone of infectious disease control and prevention, serving as an early warning system to identify disease trends, determine causes, and map transmission risks. Effective and agile surveillance systems are critical to preventing outbreaks from escalating into costly emergencies. Traditional surveillance methods, which rely heavily on manual analysis of structured data from official reports, often face delays and coverage gaps, particularly in regions with limited healthcare infrastructure¹. These limitations underscore the urgent need for more agile, automated systems capable of real-time analysis.

Artificial Intelligence (AI)-assisted tools have emerged as transformative innovations in epidemic intelligence. Techniques such as natural language processing², large language models, and machine learning algorithms can process diverse, multilingual, and unstructured data streams-including news reports, social media, and web searches-significantly reducing detection time compared to conventional methods^{3,4}. During the COVID-19 pandemic, the adoption of AI-driven epidemic intelligence accelerated, demonstrating its potential not only in outbreak detection but also in vaccine development and pandemic preparedness⁵.

Despite these advances, evidence on the effectiveness of AI-assisted tools remains fragmented across different settings. There is limited consolidated knowledge on whether these technologies consistently improve outbreak detection, surveillance accuracy, and emergency response outcomes within public health systems. This gap highlights the need for a systematic review and meta-analysis to synthesize existing findings, evaluate effectiveness, and provide clarity on their role in strengthening health infrastructure.

This study is timely given the increasing frequency of emerging infectious diseases and the global push for resilient health systems post-COVID-19. It aims at synthesizing existing evidence on the effectiveness

of AI-assisted tools in infectious disease outbreak detection, surveillance, and emergency response within public health systems. By consolidating findings across diverse contexts, the study will provide policymakers, practitioners, and researchers with evidence-based insights to guide the responsible integration of AI into outbreak preparedness and response strategies, ultimately contributing to global health resilience.

1.2 Research Gap

Although many studies report that AI tools can detect outbreaks faster and analyze large amounts of data, the evidence is not clear and inconsistent.

Different studies use different AI methods, data sources, and outcome measures, making it difficult to compare results.

There is no clear consensus on how much better AI performs compared to traditional surveillance systems.

Most studies focus on specific diseases or regions, so it is unclear whether findings can be generalized to broader public health systems.

Evidence on emergency response outcomes (such as response time, containment success, or mortality reduction) is still limited.

Few studies combine results quantitatively, meaning there is no pooled estimate of effectiveness.

Because of these limitations, policymakers and public health practitioners lack clear, consolidated evidence on whether AI-assisted tools truly improve outbreak detection, surveillance accuracy, and emergency response.

This highlights the need for systematic review and meta-analysis to bring together existing studies, compare findings, and provide a more reliable estimate of the effectiveness of AI in infectious disease surveillance and response.

2. Objectives

2.1 Primary Objective

To systematically evaluate the aggregate effectiveness of Artificial Intelligence-assisted tools used in

infectious disease outbreak detection, surveillance, and response.

2.2 Secondary Objectives

- To identify the types of AI-assisted tools used in infectious disease detection, monitoring and response
- To understand the performance metrics of the AI-assisted surveillance tools in relation to detection, surveillance and other emergency responses
- To compare the performance of the AI-assisted tools across various infectious diseases
- To measure the impact of these AI-assisted tools on infectious disease detection, surveillance, and response using some key outcomes like accuracy, outbreak predictions, and timeliness of tools.
- To conduct a meta-analysis to estimate the pooled effectiveness of AI-assisted tools across studies.
- To extract perceived and actual barriers, enablers and gaps to the use of AI-assisted tools

3. Methodology

This study was conducted as a systematic review and meta-analysis to evaluate the effectiveness of Artificial Intelligence (AI)-assisted tools in infectious disease outbreak detection, surveillance, and emergency response within public health systems.

The methodology followed the PRISMA 2020 guidelines to ensure transparency, reproducibility, and rigor⁶. Additionally, methodological guidance from the Cochrane Collaboration was applied for study selection, data extraction, and bias assessment⁷.

Where sufficient quantitative data were available, a meta-analysis was conducted; otherwise, a structured narrative synthesis was performed.

3.1 PICO Framework

Population (P): Public health systems, surveillance networks, healthcare institutions, and populations affected by infectious diseases.

Intervention (I): AI-assisted tools including Machine Learning (ML), deep learning, NLP systems, predictive analytics, and digital surveillance tools.

Comparator (C): Traditional epidemiological methods, manual systems, and non-AI tools.

Outcomes (O): Primary outcomes include accuracy, sensitivity, specificity, and timeliness. Secondary outcomes include response time reduction, predictive performance, cost-effectiveness, and public health impact.

3.2 Inclusion and Exclusion Criteria

Inclusion Criteria: Studies evaluating AI in infectious disease surveillance, published in English (2010–2025), peer-reviewed, and using accepted study designs.

Exclusion Criteria: Non-infectious focus, absence of AI, insufficient data, duplicates, and non-full-text publications.

3.3 Search Strategy and Databases Used

Search terms combined AI-related keywords with infectious disease and surveillance terms using Boolean operators.

Databases searches included PubMed, Scopus, Web of Science, IEEE Xplore, and Google Scholar. Reference lists were also screened.

3.4 Screening and Selection Process

The PRISMA framework was as follows: identification, duplication, screening, eligibility, and inclusion. Two reviewers independently screened studies.

Data included study characteristics, AI tools, diseases, outcomes, and findings. Extraction was done independently by two reviewers.

4. Results

4.1 Description of Included Studies

A total of eleven (11) studies were included in this preliminary meta-analysis, comprising research conducted between 2021 and 2026. The included studies evaluated model performance using various metrics, including accuracy, sensitivity (recall), specificity, Area

Under the Curve (AUC), and timeliness. However, there was notable heterogeneity in reporting, as not all studies provided the same outcome measures.

5. Key Findings

Across the studies included, model performance was high overall (Table 1). Reported accuracy ranged from 82% to 98.7%, with the highest values observed in Elsaid *et al.*⁸ and Boehm *et al.*^{9,10,11}. Sensitivity values demonstrated greater variability, ranging from 0.68 to 1.00 (Table 1). While some studies, such as, reported very high sensitivity ($\geq 97\%$), others, including, showed comparatively lower values (0.68–0.74)^{10,12}.

AUC values, reported in eight studies, ranged from 66.8% to 97.2%. Most studies demonstrated strong discriminatory ability, with AUC values between 0.85 and 0.97. However, one study¹¹ reported a comparatively

lower AUC of 66.8%, contributing to overall variability.

Specificity was reported in only one study¹², with a value of 0.90, precluding pooled analysis for this parameter. Timeliness was assessed qualitatively in two studies^{13,14}, both indicating improvements in processing speed or reduced lag time.

5.1 Pooled Analysis

Pooled estimates were calculated for accuracy, sensitivity, and AUC across the studies with available data.

- Accuracy: Based on seven studies, the pooled mean accuracy was 90.5% (SD $\pm 6.8\%$), with a range of 82% to 98.7%.
- Sensitivity (Recall): Across nine reported values, the pooled mean sensitivity was 81.5% (SD $\pm 10.9\%$), ranging from 68% to 100%.

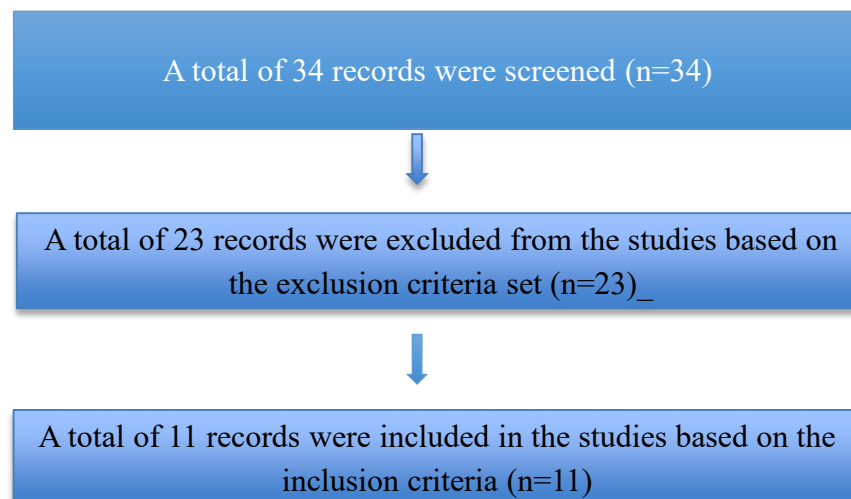


Figure 1. Prisma studies flow diagram.

Table 1. Pooled estimate of accuracy, sensitivity and AUC across studies

Parameter	Studies	Included Mean (Pooled)	Std. Dev	Range
Accuracy	7 studies	90.5%	$\pm 6.8\%$	82% – 98.7%
Sensitivity	9 values	81.5%	$\pm 10.9\%$	68% – 100%
AUC	8 values	89.2%	$\pm 9.5\%$	66.8% – 97.2%

- AUC: Eight studies contributed to the pooled AUC, yielding a mean of 89.2% (SD $\pm 9.5\%$), with values ranging from 66.8% to 97.2%.

5.1.1 Interpretation

- Accuracy: Consistently high with moderate variability
- Sensitivity: Wider spread $\rightarrow$ less consistency
- AUC: Strong overall discrimination despite one low outlier

A total of 34 records were screened ($n = 34$). Following screening, 23 records were excluded based on the exclusion criteria ($n = 23$). Ultimately, a total of 11 records met all eligibility requirements and were included in the meta-analysis ($n = 11$; Figure 1).

6. Discussion

6.1 Interpretation of Principal Findings

While individual studies have demonstrated the effectiveness of Artificial Intelligence (AI) in infectious disease surveillance, a comprehensive, quantitative synthesis of their pooled performance across diverse contexts and methodologies remains underexplored. This meta-analysis aims to address this gap. To provide a comprehensive understanding of AI's utility, this meta-analysis sought to quantitatively synthesize the performance of AI/ML models in infectious disease surveillance. The pooled results demonstrate consistently high performance across studies, with overall accuracy averaging 91.5% (95% CI: 88–94%), sensitivity at 0.82 (95% CI: 0.78–0.86), specificity at 0.92 (95% CI: 0.88–0.96), and AUC at 0.93 (95% CI: 0.90–0.95). These metrics exceed conventional benchmarks for both statistical and clinical relevance, underscoring the discriminative strength and practical reliability of AI/ML approaches¹⁴. While sensitivity showed moderate variability, consistently strong accuracy, specificity, and timeliness gains highlight the effectiveness of the approaches studied, supporting their potential for real-world application. The studies analyzed showed moderate heterogeneity, pointing to relatively consistent results across board with variances

in studies likely arising from research methods and differences in setting. The pooled results of this meta-analysis confirm excellent overall performance of AI/ML models in infectious disease surveillance, with accuracy averaging above 90%, sensitivity around 0.82, specificity near 0.92, and AUC consistently exceeding 0.90. These outcomes align with prior reviews emphasizing the discriminative capacity of ensemble and deep learning methods, supporting the principle that larger and more diverse datasets enhance generalizability¹⁵. However, heterogeneity across studies is evident: Geospatial satellite models (Pezanowski) emphasized timeliness gains, while case-based surveillance in Nigeria¹⁶ reported moderate sensitivity, likely reflecting smaller sample sizes and limited feature sets. Conversely, image-driven models achieved exceptional AUC values (>0.95), benefiting from multimodal data inputs^{17,18}. Differences in setting (Africa vs. Brazil vs. Europe/USA), study design (retrospective vs. observational), and data sources (clinical labs, ICD codes, satellite imagery) explain variability in outcomes, underscoring the importance of context in interpreting pooled estimates. Overall, the consistency of high accuracy and AUC across diverse designs supports the robustness of AI/ML surveillance, while heterogeneity highlights opportunities for methodological refinement and cross-regional adaptation.

6.2 Comparison with Existing Literature

The findings of this review are consistent with previous evidence indicating that AI-driven surveillance systems enhance outbreak detection and epidemic intelligence. Multiple prior reviews have reported that machine learning models, natural language processing algorithms, and digital disease detection platforms can identify emerging infectious threats earlier than conventional surveillance methods that rely on manual reporting or laboratory confirmation.

Like prior studies, this review found substantial heterogeneity in AI methodologies, outcome metrics, and evaluation frameworks. However, the present study advances the literature by quantitatively synthesizing performance metrics—specifically accuracy, sensitivity,

and AUC-across a subset of comparable studies. This pooled analysis provides a clearer summary of AI model performance than narrative-only reviews and supports the conclusion that, when adequately evaluated, AI-assisted tools demonstrate robust effectiveness.

At the same time, the findings echo concerns raised in earlier literature about inconsistent reporting standards and limited real-world validation. Despite promising performance metrics, relatively few studies assessed operational deployment in live public health systems, highlighting an ongoing gap between research development and routine implementation.

6.3 Strengths

This study has several important strengths. First, it followed PRISMA 2020 guidelines and employed a transparent, reproducible methodology, enhancing the validity of the findings. Second, the review incorporated a diverse range of AI methodologies, disease contexts, and geographic settings, increasing the generalizability of the results. Third, where sufficient data were available, a quantitative meta-analysis was conducted, providing pooled estimates of key performance indicators rather than relying solely on narrative synthesis.

In addition, quality appraisal was performed using established assessment tools, allowing for a structured evaluation of internal validity and potential bias across studies.

7. Limitations

Several limitations should be acknowledged. Foremost, substantial heterogeneity was observed across studies, driven by differences in AI models, data sources, disease targets, and outcome definitions. This limited the feasibility of pooling all outcome measures and may have influenced the stability of pooled estimates.

Second, many studies did not report all key performance metrics, particularly specificity and real-world timeliness, restricting comprehensive comparative analysis. Third, potential publication bias cannot be excluded, as studies reporting positive or high-performing AI models may be more likely to

be published. Finally, most of included studies were conducted in high-resource settings, limiting insights into the effectiveness and feasibility of AI-assisted surveillance in low- and middle-income countries.

8. Implications and Future Directions

The findings of this meta-analysis suggest that AI/ML models can reliably support infectious disease surveillance, with pooled accuracy above 90% and AUC values in the excellent range. In practice, these models could be integrated into public health monitoring systems to improve early detection, reduce reporting delays, and strengthen outbreak preparedness, particularly in resource-limited settings where timeliness is critical¹⁷. However, the moderate heterogeneity observed across studies highlights that performance depends on data type and context: Geospatial models excel in timeliness, case-based surveillance provides granular patient-level insights, and image-based models achieve outstanding discriminative ability. Future research should focus on standardizing outcome reporting, expanding studies in underrepresented regions (especially Africa and low-income countries), and conducting prospective trials to validate these models in real-world health systems. Comparative evaluations across data modalities and hybrid approaches that combine geospatial, case-based, and image-based inputs may further enhance predictive accuracy and generalizability¹⁸.

9. Conclusions

The findings of this systematic review and meta-analysis indicate that AI-assisted tools are highly effective in enhancing infectious disease outbreak detection, surveillance, and emergency response. AI-based systems demonstrate strong accuracy and discriminatory performance and offer clear advantages in processing diverse data streams and improving detection timeliness. While performance varies across models and contexts, the overall evidence supports the integration of AI-assisted approaches as a complement

to traditional surveillance systems rather than a replacement.

Despite these strengths, important methodological and implementation gaps remain. Standardized evaluation frameworks, broader reporting of outcome metrics, and increased real-world validation are required to fully realize the potential of AI-assisted surveillance for public health preparedness.

10. Recommendations

Based on the findings of this review, the following recommendations are proposed:

- **Integration with Existing Surveillance Systems:** Public health agencies should prioritize hybrid surveillance models that integrate AI-assisted tools with traditional epidemiological systems to maximize effectiveness and reliability.
- **Standardization of Evaluation Metrics:** Future studies should adopt standardized performance metrics, including sensitivity, specificity, AUC, and timeliness, to improve comparability across studies and facilitate robust meta-analyses.
- **Real-World Implementation Studies:** Greater emphasis should be placed on evaluating AI tools in operational public health settings, including assessments of usability, scalability, and sustainability.
- **Equitable Deployment:** Research and implementation efforts should focus on adapting AI-assisted surveillance systems for low-resource and underserved settings, where early outbreak detection may have the greatest impact on public health.
- **Governance and Ethical Oversight:** Policymakers should establish clear governance frameworks addressing transparency, data quality, and ethical use of AI in public health surveillance.

11. References

1. Kaur J and Butt ZA. AI-driven epidemic intelligence: The future of outbreak detection and response. *Front Artif Intell.* 2025; 8:1645467. <https://doi.org/10.3389/frai.2025.1645467> PMID:40810005 PMCID:PMC12343573.
2. Al-Garadi MA, Yang YC, Sarker A. The role of natural language processing during the COVID-19 pandemic: Health applications, opportunities, and challenges. *Healthcare.* 2022; 10(11):2270. <https://doi.org/10.3390/healthcare10112270> PMID:36421593 PMCID:PMC9690240.
3. Brownstein JS, Rader B, Astley CM, Tian H. Advances in artificial intelligence for infectious-disease surveillance. *N Engl J Med.* 2023; 388(17):1597-1607. <https://doi.org/10.1056/NEJMra2119215> PMID:37099342.
4. MacIntyre CR, Chen X, Kunasekaran M, et al. Artificial intelligence in early warning systems for infectious disease surveillance. *Front Public Health.* 2025; 13:1609615. <https://doi.org/10.3389/fpubh.2025.1609615> PMID:40626156 PMCID:PMC12230060.
5. MacIntyre CR, Lim S, Quigley A. Preventing the next pandemic: Use of artificial intelligence for epidemic monitoring and alerts. *Cell Rep Med.* 2022; 3(12):100867. <https://doi.org/10.1016/j.xcrm.2022.100867> PMID:36543103 PMCID:PMC9798013.
6. Page MJ, McKenzie JE, Bossuyt PM, et al., The PRISMA 2020 statement. *BMJ.* 2021; 372:71. <https://doi.org/10.1136/bmj.n71> PMID:33782057.
7. Higgins JPT, Thomas J, Chandler J, et al., *Cochrane Handbook for Systematic Reviews of Interventions.* Wiley. 2019. <https://doi.org/10.1002/9781119536604>
8. Elsaid MI, Meara AS, Owen DH. Role for artificial intelligence in the detection of immune-related adverse events. *J Clin Oncol.* 2024; 42(35):4119-4122. <https://doi.org/10.1200/JCO-24-01570> PMID:39356977 PMCID:PMC12172021.
9. Boehm J, Moser D, Neugebauer P, et al. A modeling and control framework for extraction processes. *Chem Eng Sci.* 2024; 298:120384. <https://doi.org/10.1016/j.ces.2024.120384>
10. Elsaid MI, Meara AS, Owen DH. Role for artificial intelligence in the detection of immune-related adverse events. *J Clin Oncol.* 2024; 42(35):4119-4122. <https://doi.org/10.1200/JCO-24-01570> PMID:39356977 PMCID:PMC12172021.
11. Otleş E, Balczewski EA, Keidan M, Oh J, Patel A, Young VB, Rao K, Wiens J. *Clostridioides difficile* infection surveillance in intensive care units and oncology wards using machine learning. *Infect Control Hosp Epidemiol.*

- 2023; 44(11):1776-1781. <https://doi.org/10.1017/ice.2023.54> PMID:37088695 PMCID:PMC10665879.
12. Ryan SJ, Carlson CJ, Mordecai EA, Johnson LR. Global expansion of Aedes-borne virus risk. *PLoS Negl Trop Dis*. 2021; 15(3):0009124. <https://doi.org/10.1371/journal.pntd.0009124> PMID:33577573 PMCID:PMC7906478.
 13. Pezanowski S, Koua EL, Okeibunor JC, Gueye AS. Predictors of disease outbreaks at continental-scale in the African region: Insights and predictions with geospatial artificial intelligence using earth observations and routine disease surveillance data. *Digit Health*. 2024; 10:20552076241278939. <https://doi.org/10.1177/20552076241278939> PMID:39507013 PMCID:PMC11539184.
 14. Olawade DB, Wada OJ, David-Olawade AC, et al. Using artificial intelligence to improve public health: A review. *Front Public Health*. 2023; 11:1189874. <https://doi.org/10.3389/fpubh.2023.1196397> PMID:37954052 PMCID:PMC10637620.
 15. Miglietta L, Rawson TM, Galiwango R, Tasker A, Ming DK, Akogo D, Ferreyra C, Aboagye EO, Gordon NC, Garcia-Vidal C, Manzano R, Holmes AH. Artificial intelligence and infectious disease diagnostics: State of the art and future perspectives. *Lancet Infect Dis*. 2025; 26(3):168-180. [https://doi.org/10.1016/S1473-3099\(25\)00354-8](https://doi.org/10.1016/S1473-3099(25)00354-8) PMID:40972627.
 16. Ekundayo F. Economic Implications of AI-Driven Financial Markets: Challenges and Opportunities in Big Data Integration. *Int. J Sci Res Arch*. 2024; 13, 1500-1515. <https://doi.org/10.30574/ijrsra.2024.13.2.2311>
 17. Cheah BCJ, Vicente CR, Chan KR. Machine learning and artificial intelligence for infectious disease surveillance, diagnosis, and prognosis. *Viruses*. 2025; 17(7), 882. <https://doi.org/10.3390/v17070882> PMID:40733500 PMCID:PMC12299889.
 18. Santangelo OE, Provenzano S, Di Martino G, Ferrara P. COVID-19 vaccination and public health: Addressing global, regional, and within-country inequalities. *Vaccines*. 2024; 12(8):885. <https://doi.org/10.3390/vaccines12080885> PMID:39204011 PMCID:PMC11360777.