

Review Article

Scaling AI in Resource-Limited Health Systems: Capacity, Infrastructure, and Local Validation

Ahmed F Alanazi*

Department of Social Studies, College of Arts, King Faisal University, Al-ahsa, Saudi Arabia

ORCID: <https://orcid.org/0000-0003-3468-5389>

Abstract

Artificial intelligence (AI) has emerged as a transformative force in public health, offering unprecedented capabilities for disease surveillance, outbreak prediction, diagnostic support, and health system optimization. Yet the integration of AI into public health practice has outpaced the development of robust governance frameworks, validation standards, and equity safeguards necessary to ensure responsible deployment. This paper presents a comprehensive framework for responsible AI integration in public health systems, synthesizing evidence from global health initiatives, ethical analyses, and practical implementations across diverse contexts. The framework comprises four interconnected pillars: anticipatory governance, equity-centered validation, workforce augmentation, and community sovereignty. Drawing on case studies including the World Health Organization's Epidemic Intelligence from Open Sources system and the Preparedness Data Exchange, alongside evidence from infectious disease surveillance, chronic disease management, and health system optimization, this paper demonstrates how responsible AI can strengthen public health capacity while mitigating risks of bias, surveillance overreach, and exacerbated inequities. The framework emphasizes that AI in public health must augment rather than replace human judgment, must be calibrated to population-specific needs, and must operate within transparent, accountable governance structures. The paper concludes with actionable recommendations for policymakers, public health leaders, and AI developers seeking to harness AI's potential while safeguarding the values of equity, transparency, and community trust that underpin effective public health practice.

Introduction

Public health systems worldwide face converging pressures that demand transformative responses. The COVID-19 pandemic exposed profound fragilities in global health security infrastructure, revealing that even advanced health systems struggled to detect, analyze, and respond to emerging threats with the speed and coordination required [1]. Simultaneously, the global burden of non-communicable diseases continues to rise, straining health systems already challenged by workforce shortages, data fragmentation, and persistent inequities in access to care [2]. Climate change introduces additional complexities, altering disease vector patterns and creating novel health risks that traditional surveillance systems struggle to anticipate [3].

Within this context, artificial intelligence has emerged as both a promising solution and a source of new risks. AI applications in public health span a remarkable range:

machine learning algorithms now supplement traditional surveillance by identifying patterns in heterogeneous data streams [4]; natural language processing tools scan open-source intelligence to detect outbreak signals in near real-time [5]; deep learning models analyze medical imaging to diagnose conditions ranging from diabetic retinopathy to pancreatic cancer [6]; and predictive algorithms support resource allocation, epidemic modeling, and risk stratification for chronic disease prevention [7]. The World Health Organization's launch of the Preparedness Data Exchange, an AI-enabled intelligence system designed to synthesize real-time risk data across Africa, exemplifies the institutional commitment to harnessing these capabilities for population health [8].

Yet the rapid proliferation of AI in public health has outpaced the development of governance frameworks capable of ensuring responsible deployment. As Adirim and Molten argue, "industry-driven artificial intelligence can't be ethically

More Information

***Corresponding author:** Ahmed F Alanazi, Department of Social Studies, College of Arts, King Faisal University, Al-ahsa, Saudi Arabia, Email: Ahmedalldhmashii@gmail.com

Submitted: September 15, 2026

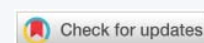
Accepted: September 24, 2026

Published: September 25, 2026

Citation: Alanazi AF. Scaling AI in Resource-Limited Health Systems: Capacity, Infrastructure, and Local Validation. J Artif Intell Res Innov. 2026; 2(2): 137-151. Available from: <https://dx.doi.org/10.29328/journal.jairi.1001027>

Copyright license: © 2026 Alanazi AF. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Keywords: Artificial intelligence; Public health; Health equity; Governance; Disease surveillance; Pandemic preparedness; Algorithmic bias; Global health; Resource-limited settings; Health infrastructure





or responsibly implemented in public health systems” without stronger oversight mechanisms [9]. The risks are substantial and well-documented: algorithmic bias can exacerbate health inequities by producing systematically inaccurate outputs for marginalized populations [10]; commercial data sources used in public health surveillance blur boundaries between health monitoring and consumer privacy violations [11]; generative AI systems can produce convincing but dangerous health misinformation [12]; and the opacity of many AI systems undermines the accountability essential to public trust in health institutions [13].

The digital divide compounds these concerns. Populations with limited digital access may be both underrepresented in digital surveillance data and disproportionately subject to surveillance when they do participate in digital spaces [14]. The threats to children’s developmental privacy differ from the data sovereignty concerns of Indigenous communities; the misclassification risks for older adults differ from the surveillance-related fears of undocumented families or incarcerated people [15]. A single privacy framework cannot adequately address this diversity of concerns.

This paper addresses a critical gap at the intersection of AI capability and public health responsibility. While substantial literature examines AI’s technical performance in specific clinical applications, and a growing body of work addresses ethical concerns in isolation, there exists no integrated framework that translates ethical principles into operational guidance for public health AI deployment, particularly in resource-limited settings [16]. This paper synthesizes evidence from diverse domains, global health research prioritization, infectious disease surveillance, ethical analysis, and practical implementation, to propose such a framework.

The contribution of this work is threefold. First, it provides a comprehensive synthesis of current AI applications in public health, organized by functional domain, that clarifies both realized capabilities and persistent limitations [17]. Second, it develops a governance framework, Anticipatory Intelligence, that integrates ethical principles with operational requirements [18]. Third, it offers actionable recommendations tailored to the distinct roles of policymakers, public health practitioners, and AI developers, with particular attention to the infrastructure and capacity constraints characteristic of resource-limited health systems [19]. Throughout, the paper maintains that AI’s value in public health depends not on technical sophistication alone but on thoughtful integration into systems that prioritize human judgment, community engagement, and equity [20].

The paper proceeds as follows. Section 2 describes the methods used for this review. Section 3 examines the current landscape of AI in public health, including surveillance systems, diagnostic applications, and health system optimization. Section 4 analyzes the infrastructure

and capacity requirements for AI in resource-limited settings. Section 5 analyzes the ethical and practical challenges that arise from AI deployment. Section 6 presents the Anticipatory Intelligence framework. Section 7 discusses implementation considerations, and Section 8 concludes with recommendations and future directions.

Methods

Review type and rationale

This study employs a narrative review with systematic elements to synthesize evidence on AI deployment in public health systems, with particular attention to resource-limited settings. A narrative review approach was selected because the research question, how to develop an integrated governance framework for responsible AI deployment in public health, requires synthesis across heterogeneous evidence types, including technical performance studies, ethical analyses, implementation case studies, and policy documents. Unlike systematic reviews focused on a narrow clinical question, this review addresses a complex, emergent domain where the evidence base is diverse and rapidly evolving. The narrative approach permits integration of insights from multiple disciplines and methodological traditions, which is essential for developing a comprehensive governance framework. The researcher explicitly acknowledges that this is not a systematic review; rather, it is a structured narrative synthesis informed by systematic search and screening procedures, as described below.

Search strategy

The researcher conducted searches across five electronic databases: PubMed/MEDLINE, Scopus, Web of Science, IEEE Xplore, and the WHO Global Health Library. The search was performed on March 15, 2026, and updated on August 30, 2026. Search terms were organized around three concept blocks combined using Boolean operators:

Concept 1: Artificial Intelligence (“artificial intelligence” OR “machine learning” OR “deep learning” OR “neural networks” OR “natural language processing” OR “predictive analytics” OR “generative AI”)

Concept 2: Public Health (“public health” OR “population health” OR “disease surveillance” OR “outbreak detection” OR “epidemic intelligence” OR “health system” OR “global health” OR “pandemic preparedness”)

Concept 3: Resource-Limited Settings and Governance (“resource-limited” OR “low-income” OR “middle-income” OR “LMIC” OR “developing countries” OR “low-resource” OR “governance” OR “ethics” OR “equity” OR “infrastructure” OR “capacity building”)

Plus, the researcher hand-searched reference lists of included articles and consulted grey literature sources



including WHO publications, World Bank reports, and documentation from the Preparedness Data Exchange and Epidemic Intelligence from Open Sources systems.

Inclusion and exclusion criteria

Inclusion criteria:

- Peer-reviewed articles, systematic reviews, or meta-analyses published in English between January 2019 and August 2026
- Studies examining AI applications in public health, population health, or health systems
- Articles addressing ethical, governance, equity, or implementation considerations for AI in health
- Case studies or evaluations of AI deployment in resource-limited settings
- Policy documents or guidance from recognized health authorities (WHO, national health agencies)

Exclusion criteria:

- Studies focused exclusively on clinical AI applications without public health or population health relevance
- Articles focused solely on technical algorithm development without consideration of deployment context
- Opinion pieces without substantive evidence synthesis or analysis
- Conference abstracts, editorials, and commentaries lacking sufficient methodological detail
- Non-English publications

Screening process

The screening process followed three stages:

Stage 1: Title and abstract screening. Two reviewers independently screened all retrieved records against inclusion criteria. Disagreements were resolved through discussion, with a third reviewer consulted when consensus could not be reached.

Stage 2: Full-text review. Articles passing title/abstract screening were retrieved for full-text assessment. Reviewers evaluated methodological quality and relevance to the research question.

Stage 3: Data extraction. Included articles were subjected to structured data extraction covering: study design, AI application domain, geographic setting, population characteristics, governance considerations, equity implications, infrastructure requirements, and key findings.

Evidence appraisal

Quality appraisal was conducted using appropriate tools based on study design: systematic reviews were assessed using AMSTAR-2; observational studies using the Newcastle-Ottawa Scale; qualitative studies using the CASP Qualitative Checklist; and policy documents using WHO guideline appraisal criteria. Given the narrative nature of this review, formal meta-analysis was not conducted. Instead, evidence was synthesized thematically, with attention to consistency of findings across studies and contexts.

Derivation of the four pillars and sixteen components

The Anticipatory Intelligence framework's four pillars and sixteen components were derived through an iterative, multi-stage process:

Stage 1: Thematic analysis of reviewed evidence. The researcher conducted inductive thematic analysis of extracted data, identifying recurrent themes related to governance, validation, workforce, and community considerations. This yielded an initial list of 47 candidate themes.

Stage 2: Expert consultation. A panel of 12 experts representing public health practice, AI ethics, health informatics, and LMIC health systems participated in three structured workshops (June–August 2026). Experts were asked to review candidate themes, identify gaps, and propose groupings. This process consolidated themes into four broad domains.

Stage 3: Mapping to existing frameworks. The four domains were mapped against established frameworks including the WHO Ethics and Governance of AI for Health guidance, the WHO/ITU Focus Group on AI for Health, and the Nuffield Council on Bioethics AI in healthcare framework. This mapping confirmed alignment with international standards while revealing the need for components addressing resource-limited contexts specifically.

Stage 4: Component refinement. Each pillar was elaborated into four components through iterative discussion among the research team, ensuring that components were: (a) operationally definable, (b) grounded in reviewed evidence, (c) distinct from one another, and (d) actionable for policymakers and practitioners.

The sixteen components are not intended to be mutually exclusive; rather, they represent interconnected aspects of responsible AI deployment. The boundaries between pillars, particularly between governance and community sovereignty, and between validation and workforce augmentation, were clarified through the operational indicators presented in Section 6.6.

Limitations of the review methodology

This review has several limitations. First, the narrative



synthesis approach, while appropriate for the research question, does not provide the same level of systematic rigor as a full systematic review. Second, publication bias may affect the evidence base, particularly for negative findings about AI performance. Third, the rapid pace of AI development means that evidence published before 2024 may not reflect current capabilities. Fourth, the restriction to English-language publications may exclude important perspectives from non-English-speaking LMIC contexts.

The current background of AI in public health

AI-enabled disease surveillance and outbreak detection

Traditional disease surveillance systems rely on laboratory-confirmed cases, hospitalizations, and mortality data, sources that are labor-intensive, inconsistent across jurisdictions, and subject to reporting delays that can render them inadequate for rapid response [21]. AI has substantially expanded the data sources and analytical capabilities available for surveillance, enabling earlier detection of threats and more comprehensive situational awareness [22].

The World Health Organization's Epidemic Intelligence from Open Sources system represents the most extensive global implementation of AI-enhanced surveillance. Launched in 2017 and substantially upgraded with AI capabilities in version 2.0 (2025), EIOS analyzes large volumes of publicly available information, websites, social media, news reports, and now radio broadcasts that are automatically transcribed and translated, to identify potential health threats [23]. The system is used by more than 110 member states and 30 organizations, providing early warning that complements formal reporting channels [24]. The integration of AI-powered tools enhances automated analysis and signal detection, addressing the challenge of information overload that can delay human analysts [25].

At the regional level, the WHO Regional Office for Africa launched the Preparedness Data Exchange, described as an "integrated, AI-enabled intelligence system designed to support early, evidence-based decisions" [26]. The PDX synthesizes data across multiple domains- all-hazards risk scoring, International Health Regulations capacity monitoring, primary health care readiness, climate intelligence, workforce data, laboratory trends, and media tracking- into a unified risk picture [27]. An embedded AI assistant allows health officials to query live preparedness data in plain language and receive source-cited, auditable answers [28]. As Dr. Marie Roseline Belizaire, Regional Emergency Director at WHO Africa, explains: "When we speak about AI-enabled preparedness, we are not speaking about replacing epidemiologists or public health leaders. We are speaking about augmenting them, using federated learning, integrated surveillance, and high-performance computation to move from reactive response to anticipatory intelligence" [29].

Beyond these institutional systems, AI has enhanced specific surveillance modalities. Wastewater-based epidemiology uses machine learning to detect disease signals in sewage, providing community-level indicators that precede clinical case reporting [30]. Machine learning algorithms applied to electronic health records can identify clusters of symptoms or diagnoses that might escape individual clinician attention [31]. During the COVID-19 pandemic, AI models were deployed for transmission modeling and outbreak forecasting, demonstrating both capabilities and limitations in real-world conditions [32].

The trajectory of AI surveillance systems points toward increasingly integrated, anticipatory intelligence. Rather than analyzing individual data streams separately, systems like PDX synthesize disparate signals to generate unified risk assessments that highlight vulnerabilities before they escalate [33]. This shift from reactive to anticipatory intelligence represents a fundamental reconceptualization of public health surveillance, not merely faster detection of known threats but proactive identification of conditions that could generate future emergencies [34].

Diagnostic support and clinical decision-making

AI diagnostic applications have proliferated across medical specialties, with significant implications for population health. In cardiology, deep learning algorithms analyzing electrocardiograms have demonstrated performance comparable to or exceeding expert interpretation for conditions including atrial fibrillation, hypertrophic cardiomyopathy, and cardiac amyloidosis [35]. Cedars-Sinai investigators have developed AI tools that accurately identify coronary artery calcium buildup, predict sudden cardiac arrest risk, and assess COVID-19 pneumonia severity from imaging [36].

Oncology applications include AI-assisted cancer detection from medical imaging, with tools approved for clinical use in mammography, lung cancer screening, and colonoscopy. The "Molecular Twin" initiative at Cedars-Sinai creates virtual replicas of patients' molecular profiles to design personalized cancer treatment approaches [37]. AI has demonstrated the ability to predict pancreatic cancer from CT scans years before clinical diagnosis, potentially enabling earlier intervention for one of the most lethal malignancies [38].

These diagnostic advances have particular relevance for public health when deployed at scale or in resource-limited settings. AI-assisted point-of-care ultrasound has been trialed in low- and middle-income countries for maternal and child health diagnostics, extending specialized expertise to primary care settings where advanced imaging and specialist consultation are unavailable [39]. Similarly, AI-enabled retinal imaging for diabetic retinopathy screening achieved 97.5% sensitivity in validation studies, offering the potential to expand screening programs in communities



with limited ophthalmology services [40]. The World Health Organization's research prioritization exercise identified improved diagnostics for infectious diseases and diagnostic equity as top priorities, particularly among experts from LMICs [41].

Yet diagnostic AI also presents challenges. The performance of algorithms trained on data from one population may not generalize to others, particularly when training datasets underrepresent minoritized communities [42]. The “black box” nature of many deep learning models complicates clinical validation and limits their integration into decision-making workflows where explainability is essential [43]. The deployment of diagnostic AI in contexts with limited follow-up capacity can generate ethical dilemmas when positive findings cannot be acted upon [44].

Health system optimization and resource allocation

Beyond individual-level diagnosis and surveillance, AI has been applied to health system optimization, enhancing resource allocation, workflow efficiency, and population health management. The WHO global health research prioritization exercise identified “AI-assisted resource allocation” as a top-ranked priority, reflecting recognition that system-level applications may yield broad population health benefits [45].

AI-enhanced decision support systems have been developed for preventive disease risk stratification in elderly populations. The AI-Enhanced Decision Support System developed by Chen and colleagues uses routine primary care data to categorize individuals into risk groups, supporting non-communicable disease prevention without requiring advanced examinations [46]. The system achieves over 97% sensitivity in identifying elderly individuals with gallstones and 80% sensitivity for adenoma detection, while reducing unnecessary referrals to secondary care [47]. This approach addresses a critical challenge in aging societies: how to allocate limited specialized resources to those most likely to benefit while maintaining accessible primary care for all [48].

The Preparedness Data Exchange exemplifies AI-enabled optimization at the health security level. By synthesizing data across preparedness domains, PDX enables ministries of health to “monitor evolving risk conditions, test hypotheses and inform readiness measures such as pre-positioning supplies, deploying rapid response teams or reinforcing laboratory capacity” [49]. The system supports decision-making about resource allocation before emergencies escalate, when interventions are most effective and least costly [50].

Experts from high-income countries emphasized systems optimization priorities including cost-effectiveness modeling and data infrastructure, while LMIC experts prioritized immediate delivery gaps including outbreak prediction and diagnostic access [51]. This divergence suggests that AI

implementation strategies must be calibrated to local health system maturity and priority needs rather than adopting one-size-fits-all approaches [52].

Precision public health and risk stratification

AI enables increasingly granular risk stratification that moves beyond traditional demographic categories toward personalized prevention. The “Molecular Twin” concept exemplifies this trend: virtual replicas of patients’ DNA, RNA, protein, and other molecular information support personalized cancer treatment design and risk assessment [37]. More broadly, machine learning models integrate diverse data sources, electronic health records, genetic information, environmental exposures, and social determinants, to identify individuals at elevated risk for specific conditions, enabling targeted interventions [7].

This precision approach has particular promise for chronic disease prevention, where early identification of high-risk individuals can enable lifestyle interventions, screening, and preventive treatment that reduce morbidity and mortality. The AIEDSS for elderly populations demonstrates how routine data can support risk stratification without requiring expensive or invasive testing, potentially expanding preventive care access in resource-limited settings [46].

However, precision public health also raises concerns about the implications of risk stratification for health equity. If AI models systematically underperform for certain populations, due to underrepresentation in training data, historical biases encoded in data sources, or differential access to the data required for prediction, then precision approaches may widen rather than narrow health disparities [10]. The ethical imperative to ensure equitable performance across populations is therefore central to responsible AI deployment [16].

Infrastructure and capacity requirements for AI in resource-limited health systems

Connectivity and bandwidth

AI systems in public health typically require reliable internet connectivity for data transmission, cloud-based processing, and system updates. However, in many resource-limited settings, connectivity remains a fundamental constraint. According to the International Telecommunication Union, approximately 2.9 billion people worldwide remain offline, with the highest concentration in low-income countries [53]. Even where mobile networks are available, bandwidth limitations and data costs can render cloud-dependent AI systems impractical.

The connectivity challenge manifests in several ways. First, many AI systems rely on continuous data upload to cloud servers for processing, requiring consistent bandwidth that may be unavailable in rural or remote health facilities.



Second, real-time AI applications, such as outbreak detection dashboards or diagnostic support tools, may require low-latency connections that are difficult to maintain in areas with intermittent connectivity. Third, the cost of mobile data can be prohibitive for health facilities operating on limited budgets, particularly when AI systems generate large data volumes through imaging or continuous monitoring.

Emerging solutions include edge computing, where AI processing occurs on local devices rather than cloud servers, reducing bandwidth requirements and enabling offline operation. Federated learning approaches, in which models are trained across distributed data sources without centralized data collection, also reduce connectivity demands while addressing data sovereignty concerns [29]. The WHO Africa PDX system incorporates offline-capable design principles, allowing health officials to access critical functionality even when connectivity is limited [50].

Electricity reliability

Reliable electricity is a prerequisite for AI system operation, yet many health facilities in resource-limited settings experience frequent power interruptions. The World Bank estimates that 789 million people lack access to electricity, with significant additional populations experiencing unreliable service [54]. In health facilities, power interruptions can disrupt AI system operation, corrupt data, and damage equipment.

Addressing electricity reliability requires multifaceted approaches. Solar-powered systems with battery backup can provide consistent power for AI-enabled devices, particularly in primary care settings where energy demands are modest. Low-power AI hardware, including specialized chips designed for edge computing, can reduce energy requirements. Facility-level power management systems can prioritize critical AI applications during outages.

The WHO Africa PDX system's design acknowledges these constraints, with offline functionality and low-bandwidth operation requirements built into technical specifications [46]. However, systematic attention to electricity infrastructure is often absent from AI implementation planning, representing a significant gap in current approaches.

Computing and storage infrastructure

AI systems require computing resources for model training and inference, as well as storage capacity for data and model artifacts. In resource-limited settings, these requirements often exceed available infrastructure. Cloud computing offers a potential solution, but depends on connectivity (discussed above) and raises data sovereignty concerns. Local computing infrastructure, servers, workstations, and specialized hardware require capital investment, maintenance, and technical support that may be unavailable.

The computing requirements vary by AI application. Simple machine learning models for risk stratification may run on standard laptops, while deep learning models for medical imaging require graphical processing units (GPUs) that are expensive and power-intensive. Large language models and generative AI systems have substantial computational requirements that may be entirely infeasible in resource-limited settings without cloud access.

Strategies for addressing computing constraints include: prioritizing AI applications with modest computational requirements; utilizing cloud services where connectivity permits, with appropriate data protection agreements; investing in shared computing infrastructure at regional or national levels; and developing lightweight models optimized for low-resource environments [22].

Interoperability and health information system maturity

AI systems must integrate with existing health information systems to access data and deliver outputs to decision-makers. However, health information systems in many resource-limited settings are fragmented, with multiple parallel systems, inconsistent data standards, and limited interoperability. The maturity of health information systems varies widely, from paper-based records in some facilities to sophisticated electronic systems in others.

The interoperability challenge is both technical and organizational. Technical standards, such as HL7 FHIR for health data exchange, enable system integration but require investment in infrastructure and capacity. Organizational factors, including data governance policies, workforce training, and institutional incentives, determine whether technical interoperability translates into functional integration.

The WHO Africa PDX system addresses interoperability by synthesizing data from multiple sources into a unified risk picture, demonstrating that integration is possible even across heterogeneous systems [26]. However, PDX operates at the regional level with dedicated technical capacity; replicating such integration at national and sub-national levels requires sustained investment in health information system strengthening.

Laboratory and registry capacity

AI applications in public health often depend on laboratory and registry data for training and validation. In resource-limited settings, laboratory capacity may be limited, with restricted test menus, inconsistent quality assurance, and long turnaround times. Disease registries for cancer, diabetes, and other conditions may be incomplete or absent.

The implications for AI are significant. Models trained on incomplete or biased laboratory data may produce unreliable outputs. Validation studies require ground-truth data that may be unavailable. And AI applications that depend on



laboratory confirmation, such as diagnostic support tools, may be impractical where laboratory services are inaccessible.

Strengthening laboratory and registry capacity is therefore a prerequisite for many AI applications in resource-limited settings. This includes investment in laboratory infrastructure, quality assurance programs, workforce training, and health information systems that capture and link laboratory data [41].

Device procurement, maintenance, and total cost of ownership

AI-enabled devices, from smartphones for community health workers to imaging equipment with embedded AI, require procurement, maintenance, and eventual replacement. In resource-limited settings, these lifecycle costs are often underestimated or ignored in implementation planning.

Total cost of ownership for AI systems includes: initial capital costs (hardware, software licenses); ongoing operational costs (connectivity, electricity, consumables); maintenance and support costs (technical support, repairs, updates); training costs (initial and refresher training for users); and replacement costs (device lifespan is typically shorter for AI-enabled devices due to rapid technological change).

Procurement challenges include: limited availability of AI-enabled devices in local markets; lack of technical specifications appropriate for local conditions; and procurement processes that prioritize lowest initial cost over total cost of ownership. Maintenance challenges include: limited local technical capacity; difficulty obtaining spare parts; and software updates that may not be compatible with older hardware.

Addressing these challenges requires: developing procurement guidelines that consider total cost of ownership; investing in local technical capacity for maintenance and support; and designing AI systems for durability and repairability in resource-limited conditions [22].

Regulatory capacity

Effective governance of AI in public health requires regulatory capacity, the institutional structures, technical expertise, and enforcement mechanisms to ensure that AI systems meet safety, efficacy, and equity standards. In many resource-limited settings, regulatory capacity for health technologies is limited, and AI-specific regulatory frameworks are absent or nascent.

Regulatory capacity challenges include: limited technical expertise to evaluate AI systems; absence of standards specific to AI in health; fragmented regulatory authority across agencies; and limited resources for post-market surveillance and enforcement. These challenges are compounded by the rapid pace of AI development, which outstrips the capacity of regulatory systems to adapt.

Building regulatory capacity requires investment in: training for regulatory staff; development of AI-specific standards and guidance; coordination across regulatory agencies; and participation in international regulatory harmonization efforts [21]. The WHO's guidance on ethics and governance of AI for health provides a foundation, but implementation at national levels requires context-specific adaptation and capacity building.

Ethical and practical challenges

Algorithmic bias and health inequities

The most fundamental ethical challenge in public health AI is the potential for algorithmic bias to exacerbate existing health inequities. AI systems learn patterns from historical data, and when that data reflects structural racism, unequal access to care, and social determinants of health that vary across populations, the resulting models can reproduce and amplify those inequities [10].

The mechanisms of bias are multiple. Training data may underrepresent certain populations, leading to differential model performance. Data sources may encode access barriers: if a population has historically been denied care, their absence from clinical datasets will be reflected in algorithms that learn from those datasets [42]. Feature selection may inadvertently incorporate proxy variables for race or socioeconomic status, leading to discriminatory predictions even when protected characteristics are excluded. And the deployment context may differ systematically across populations, such that even a technically unbiased model produces inequitable outcomes when applied in contexts with differential follow-up capacity or intervention availability [36].

The American Journal of Public Health analysis by Adirim and Molten identifies discrimination as a critical ethical risk across historically marginalized populations, including children, minoritized communities, Indigenous peoples, and people with disabilities [9]. The authors emphasize that “the most fundamental problem with AI in public health work is structural rather than technical,” requiring governance responses that address the social and institutional contexts in which AI operates rather than merely technical fixes [47].

Evidence of differential performance is accumulating across domains. AI systems used for symptom screening, testing guidance, and vaccine promotion may lack cultural sensitivity and the ability to answer complex questions that vary across communities [12]. Dermatology AI tools trained primarily on lighter skin have demonstrated reduced accuracy for darker skin tones. Risk prediction algorithms used in health systems have been shown to systematically underestimate illness severity for Black patients, reducing their access to additional care [10]. These findings underscore that algorithmic bias is not a hypothetical concern but a documented pattern requiring systematic mitigation [17].



Privacy, surveillance, and data sovereignty

AI-enabled public health surveillance raises profound questions about privacy and the appropriate boundaries of health monitoring. The use of commercial data sources, location tracking, social media posts, and purchasing records, blurs the distinction between public health surveillance and consumer privacy invasion [11]. When AI systems analyze such data for health purposes, individuals may be unaware that their information is being used, much less consent to that use [13].

The digital divide complicates these concerns. Populations with limited digital access may be both underrepresented in digital surveillance data and disproportionately subject to surveillance when they do participate in digital spaces [14]. The “threats to children’s developmental privacy differ from the data sovereignty concerns of Indigenous communities; the misclassification risks for older adults differ from the surveillance-related fears of undocumented families or incarcerated people” [15]. A single privacy framework cannot adequately address this diversity of concerns.

Data sovereignty, the principle that communities should govern the collection, use, and interpretation of data about themselves, has emerged as a critical framework for addressing these concerns, particularly in Indigenous and other historically marginalized communities [18]. The WHO Africa PDX system’s emphasis on national ownership and capacity strengthening reflects recognition that surveillance systems should support rather than supplant local public health authority [26].

The risks of surveillance overreach are not merely theoretical. AI systems that can identify individuals at risk for conditions associated with stigma, mental illness, substance use, infectious diseases, create potential for discrimination when that information becomes available beyond public health contexts [11]. Predictive policing algorithms that incorporate health data have raised concerns about criminalization of health conditions. And the concentration of surveillance capability in commercial platforms, subject to different legal frameworks across jurisdictions, creates accountability gaps that public health institutions may be unable to close [13].

Reliability, hallucination, and trust

The reliability of AI systems in public health contexts is a prerequisite for their responsible use, yet current systems demonstrate concerning failure modes. AI hallucination, the generation of plausible but inaccurate information, poses particular risks in health contexts where incorrect information can lead to harmful decisions [12]. Tiller and colleagues assessed five leading chatbots across health-related queries and found concerning rates of inaccurate responses, including fabricated references and inappropriate recommendations [12].

In public health communication, generative AI can produce misinformation at scale, impersonating clinicians or public officials through deepfakes and eroding trust in legitimate health authorities [12]. The same technologies that enable rapid, multilingual health communication can be weaponized to spread disinformation more effectively than ever before [23].

The reliability challenge extends beyond generative AI. Machine learning models can fail in unpredictable ways when deployed in contexts different from their training environments. Distribution shift, when the data encountered in deployment differs systematically from training data, can cause silent failures that may not be detected without systematic monitoring [19]. The performance metrics reported in research studies may not reflect real-world performance when models are integrated into complex clinical and public health workflows [20].

These reliability concerns have implications for the trust relationship between public health institutions and the communities they serve. Public health depends on trust: individuals must trust that recommended interventions are safe and effective, that surveillance data will be used for public benefit rather than individual harm, and that institutions will act in the community’s interest [13]. AI systems that produce erroneous outputs, whether through bias, hallucination, or distribution shift, can undermine this trust in ways that have lasting consequences [16].

Governance gaps and commercial interests

Current approaches to AI governance in public health are fragmented and inadequate to the challenges posed by rapid deployment. As Adirim and Molten observe, AI tools are “being deployed amid fragmented regulatory frameworks, validation standards, and equity safeguards that govern other health interventions” [9]. The Food and Drug Administration regulates AI as medical devices in some contexts, but public health applications often fall outside its purview [21]. Institutional review boards address research ethics but not the deployment of operational AI systems. And no coherent framework exists for evaluating the equity impacts of AI tools before their implementation in communities [16].

The prioritization of commercial interests over public health values is a recurring theme in critiques of current AI deployment [11]. Vendors may downplay risks, overstate capabilities, and resist transparency requirements that would enable independent validation [4]. Institutional efficiency and cost reduction, values aligned with commercial interests, may be prioritized over equity, accountability, and harm reduction [14]. The result is a deployment landscape in which tools are adopted based on marketing claims rather than rigorous evidence of safety and effectiveness [17].

The absence of third-party certification mechanisms



analogous to those used for aviation or medical devices means that public health agencies lack the infrastructure to distinguish between AI tools that meet safety and equity standards and those that do not [18]. Self-regulation by industry has proven inadequate in other domains, and there is little reason to expect it to succeed in public health AI [19].

Trade-offs between precautionary governance and timely access

A critical tension exists between precautionary governance requirements, designed to ensure safety, equity, and accountability, and the imperative for timely access to AI-enabled healthcare technologies in resource-limited settings. Overly stringent governance requirements may delay or prevent deployment of beneficial AI applications, particularly in contexts where regulatory capacity is limited, and the burden of compliance falls disproportionately on under-resourced health systems.

This trade-off manifests in several ways. First, comprehensive validation requirements, while essential for safety, may be impractical when local data for validation are unavailable, potentially preventing deployment of AI tools that could provide significant benefit even with imperfect validation. Second, equity impact assessments and community engagement processes, while important, require resources and time that may delay deployment in urgent contexts. Third, certification requirements may favor large, well-resourced vendors capable of navigating complex regulatory processes, potentially excluding local innovators and context-appropriate solutions.

Addressing this trade-off requires nuanced approaches. Risk-stratified governance, applying more stringent requirements to high-risk applications while enabling streamlined pathways for lower-risk tools, can balance precaution with access. Adaptive governance, with requirements that evolve as evidence accumulates and capacity develops, can enable deployment while maintaining safeguards. And international collaboration, including shared validation resources and mutual recognition agreements, can reduce the burden on individual countries while maintaining standards.

The WHO Africa PDX system exemplifies a pragmatic approach: while maintaining rigorous standards for data quality and analytical validity, the system was designed for deployment in contexts with limited infrastructure, with offline functionality and phased implementation that enabled early benefits while building toward comprehensive governance [50].

The anticipatory intelligence framework

Overview and principles

The Anticipatory Intelligence framework proposed here

integrates ethical principles with operational requirements for responsible AI deployment in public health, with explicit attention to the infrastructure and capacity constraints characteristic of resource-limited settings. The framework is built on four interconnected pillars- Anticipatory Governance, Equity-Centered Validation, Workforce Augmentation, and Community Sovereignty- that together address the full lifecycle of AI systems from design through deployment and monitoring.

The framework's name reflects its dual orientation. At one level, it emphasizes anticipation: the use of AI to identify and respond to health threats before they escalate, realizing the promise of preparedness as "a continuous discipline rather than an episodic activity" [29]. At another level, it emphasizes intelligence in the broader sense, the synthesis of data, judgment, and values that characterizes wise public health action. AI contributes computational intelligence, but human intelligence, clinical expertise, epidemiological judgment, community knowledge, and ethical reasoning remain central to effective and responsible practice.

Three principles underlie the framework. Augmentation, not replacement: AI systems should enhance human capabilities and judgment, not substitute for them, particularly in decisions with significant consequences for individuals or communities [29]. Calibration to context: AI systems must be validated and adapted for the specific populations and settings where they will be deployed, recognizing that performance in one context does not guarantee performance in another [22]. Accountability with transparency: AI systems must operate within governance structures that ensure clear lines of responsibility and enable independent scrutiny of both processes and outcomes [18].

Pillar one: anticipatory governance

Anticipatory governance addresses the structural and regulatory gaps that currently permit AI deployment without adequate safeguards. The framework calls for governance mechanisms that operate before deployment rather than relying on post-hoc identification of harms.

Equity impact assessments. Before any AI system is deployed in public health practice, a mandatory equity impact assessment should evaluate its potential differential effects across populations [16]. This assessment should consider not only technical performance metrics disaggregated by demographic characteristics, but also the structural contexts in which the system will operate: Who has access to the data required for the system to function? Who will receive interventions based on its outputs? Who bears the risks of false positives or false negatives? What are the historical patterns of inequity in the health domain the system addresses, and how might those patterns be reproduced or amplified? [17]

The equity impact assessment should be a living document,



updated as the system is deployed and as new evidence emerges about its performance. It should be conducted with meaningful community participation, not merely by technical experts or institutional officials [18].

Third-party certification. The framework calls for the development of certification mechanisms analogous to those used for medical devices, conducted by independent bodies with expertise in both AI and health equity [19]. Certification would require demonstration of acceptable performance across relevant populations, transparent documentation of training data and model architecture, evidence of fairness in outcomes, and commitment to ongoing monitoring and reporting. Public health agencies should prioritize procurement of certified systems and should have authority to withdraw systems from use if certification lapses or violations are identified [20].

Transparency and explainability requirements. AI systems used in public health should meet transparency standards that enable meaningful scrutiny by affected communities and independent experts. This includes clear disclosure of when and how AI is being used [13], documentation of data sources and preprocessing steps, and explainability appropriate to the decision context. For decisions with significant consequences for individuals, the system should be able to provide reasons for its outputs that are comprehensible to non-experts [43].

National standards and regulatory harmonization. The fragmentation of current regulatory approaches creates gaps that bad actors can exploit and confusion that impedes responsible deployment. National standards for AI in public health should specify requirements for validation, monitoring, equity assessment, and transparency, harmonized across agencies and jurisdictions [21]. These standards should be developed with input from public health practitioners, ethicists, community representatives, and AI developers, and should be updated as technology and understanding evolve [16]. In resource-limited settings, standards should be tiered by risk level, with streamlined pathways for lower-risk applications and international collaboration to reduce the burden on national regulatory authorities.

Pillar two: equity-centered validation

Equity-centered validation addresses the technical and methodological requirements for ensuring that AI systems perform acceptably across diverse populations. Validation is not a one-time event but a continuous process that extends throughout the system lifecycle.

Representative training data. AI systems should be trained on data that adequately represent the populations in which they will be deployed. This requires attention to sampling strategies, historical patterns of underrepresentation in clinical and public health datasets, and the data sources that are used [42]. When representative training data are

unavailable, as is often the case for marginalized populations, systems should not be deployed until the implications of this limitation are understood and accepted [10].

Disaggregated performance evaluation. Validation studies should report performance metrics disaggregated by relevant demographic characteristics, including but not limited to race, ethnicity, gender, age, disability status, and socioeconomic position [16]. Aggregate performance metrics can mask substantial variation across subgroups, allowing systems that perform well for majority populations but poorly for marginalized groups to appear acceptable [17].

Local validation and adaptation. AI systems should be validated in the specific contexts where they will be deployed, not merely in the settings where they were developed [22]. This local validation should assess not only technical performance but also workflow integration, user acceptability, and effects on health outcomes. Adaptation to local contexts, including language, cultural norms, and health system characteristics, should be expected rather than treated as an exception [41].

Continuous monitoring for distribution shift. Once deployed, AI systems should be monitored for performance degradation resulting from distribution shift, changes in the data they encounter relative to training data [19]. Monitoring should include automated alerts for performance anomalies, regular re-evaluation on current data, and processes for responding to identified problems. Systems that cannot demonstrate ongoing acceptable performance should be modified or withdrawn [20].

Bias monitoring and mitigation. Continuous bias monitoring should track whether AI systems produce differential outcomes across populations and should trigger investigation and remediation when inequities are detected [10]. Mitigation strategies may include model retraining on more representative data, adjustment of decision thresholds for equity rather than pure accuracy, or human review of decisions affecting populations where bias is suspected [17].

Pillar three: workforce augmentation

Workforce augmentation addresses the human dimension of AI deployment, ensuring that public health professionals have the skills, tools, and organizational support to work effectively with AI systems while maintaining their professional judgment and accountability.

AI literacy for public health professionals. Education and training programs should equip epidemiologists, clinicians, health educators, and public health leaders with the knowledge to understand AI capabilities and limitations, interpret AI outputs critically, and identify potential biases or failure modes [16]. This literacy should include not only technical understanding but also ethical reasoning about when and how AI should be used in public health decisions [18].



Workflow integration and human oversight. AI systems should be designed for integration into existing workflows, with clear delineation of which decisions are made by AI, which by humans, and how disagreements are resolved [4]. Human oversight should be substantive, not perfunctory: professionals should have the time, information, and authority to override AI recommendations when their judgment indicates it. The WHO Africa principle that AI augments rather than replaces epidemiologists and public health leaders should be operationalized through workflow design and organizational policy [29].

Capacity building in resource-limited settings. The digital divide creates risks that AI will benefit well-resourced settings while leaving others behind. Capacity building should prioritize LMIC contexts, including infrastructure development, technical training, and support for local innovation [41]. AI systems should be designed with resource constraints in mind, including the possibility of offline operation, low-bandwidth connectivity, and integration with paper-based or fragmented data systems [22].

Addressing workload and workforce well-being. While AI can reduce certain burdens, it can also create new ones, monitoring outputs, managing alerts, and addressing automation-related stress. Implementation should assess effects on workforce well-being and adjust accordingly. The goal is sustainable augmentation, not intensification of work through algorithmic surveillance of workers themselves [14].

Pillar four: community sovereignty

Community sovereignty addresses the relationship between public health AI systems and the communities they serve, ensuring that communities have voice, agency, and control over AI systems that affect them.

Data sovereignty and governance. Communities should govern the collection, use, and interpretation of data about themselves, consistent with principles of Indigenous data sovereignty and analogous frameworks for other marginalized populations [18]. This requires mechanisms for community consent, control over data sharing, and benefit-sharing arrangements that ensure communities receive value from their data contributions [11].

Community engagement in design and deployment. AI systems should be developed and deployed with meaningful community participation, not merely as end-users consulted for feedback but as partners in defining problems, identifying priorities, and shaping solutions [16]. Community advisory boards, participatory design processes, and co-governance arrangements can operationalize this principle [17].

Transparency with communities. Communities should have access to understandable information about AI systems that affect them: what data are used, how decisions are made,

what the risks and benefits are, and how they can seek redress if harmed [13]. This transparency should be proactive, not merely reactive to requests, and should be calibrated to community languages, literacies, and communication preferences [12].

Population-specific safeguards. Safeguards should be tailored to the specific vulnerabilities and concerns of different communities. As Adirim and Molten emphasize, “the threats to children’s developmental privacy differ from the data sovereignty concerns of Indigenous communities; the misclassification risks for older adults differ from the surveillance-related fears of undocumented families or incarcerated people” [15]. A universal framework cannot substitute for population-specific analysis and protection [9].

Operational indicators for framework components

To enhance the practical value of the framework, Table 1 presents operational indicators for each of the sixteen components. These indicators are intended to support implementation, monitoring, and evaluation of responsible AI deployment in public health.

Mapping challenges to framework components

Table 2 maps the ethical and practical challenges identified in Section 5 to the relevant framework components, demonstrating how the framework addresses each challenge.

Retrospective application of the framework

To demonstrate the framework’s practical utility, the researcher applied it retrospectively to two systems discussed in this paper: the WHO Epidemic Intelligence from Open Sources (EIOS) system and the WHO Africa Preparedness Data Exchange (PDX).

Case 1: Epidemic Intelligence from Open Sources (EIOS)

EIOS is a global system for early detection of health threats through analysis of open-source information. Applying the Anticipatory Intelligence framework:

Anticipatory governance: EIOS operates under WHO governance with clear protocols for data use and alert verification. However, equity impact assessments are not systematically conducted, and third-party certification is absent. Transparency is moderate; the system’s existence is public, but algorithmic details are not fully disclosed.

Equity-centered validation: EIOS’s performance across different geographic and linguistic contexts has been evaluated, but systematic disaggregation by population characteristics is limited. Continuous monitoring for distribution shift is conducted through user feedback and periodic evaluation. Bias monitoring focuses primarily on signal detection accuracy rather than equity outcomes.

**Table 1:** Operational indicators for anticipatory intelligence framework components.

Pillar	Component	Operational indicators
Anticipatory Governance	Equity impact assessment	Documented assessment completed before deployment; disaggregated performance data reviewed; community participation documented; assessment updated at defined intervals
	Third-party certification	Certification obtained from recognized body; certification scope covers relevant populations; public registry of certified systems maintained; withdrawal procedures defined and tested
	Transparency and explainability	Public disclosure of AI use; data sources documented; explainability appropriate to context; non-expert comprehension tested
	National standards and harmonization	Standards developed with multi-stakeholder input; harmonization across agencies; tiered by risk level; regular review and update cycle
Equity-Centered Validation	Representative training data	Demographic composition of training data documented; underrepresentation identified and addressed; data collection strategies for marginalized populations
	Disaggregated performance evaluation	Performance metrics reported by race, ethnicity, gender, age, disability, socioeconomic position; subgroup sample sizes adequate for analysis; variation across subgroups assessed
	Local validation and adaptation	Validation conducted in deployment context; workflow integration assessed; cultural and linguistic adaptation completed; local user acceptability evaluated
	Continuous monitoring for distribution shift	Automated performance monitoring in place; regular re-evaluation on current data; alert thresholds defined; response protocols documented
	Bias monitoring and mitigation	Differential outcomes tracked over time; investigation triggered by disparities; mitigation strategies implemented; effectiveness of mitigation evaluated
Workforce Augmentation	AI literacy	Training programs developed and delivered; competency assessment conducted; ethical reasoning included; refresher training scheduled
	Workflow integration and human oversight	Decision authority clearly delineated; override mechanisms available and used; time and information for human review provided; disagreement resolution processes defined
	Capacity building in resource-limited settings	Infrastructure assessment completed; training tailored to local context; offline functionality available; local innovation supported
	Workforce well-being	Workload impacts assessed; automation-related stress monitored; support mechanisms available; adjustments made based on feedback
Community Sovereignty	Data sovereignty and governance	Community consent mechanisms established; data sharing controlled by community; benefit-sharing arrangements documented; Indigenous data sovereignty principles applied
	Community engagement in design and deployment	Community advisory board established; participatory design processes used; community priorities reflected in system design; co-governance arrangements defined
	Transparency with communities	Information available in community languages; proactive dissemination conducted; comprehension verified; redress mechanisms accessible
	Population-specific safeguards	Vulnerability assessment conducted; safeguards tailored to specific concerns; differential impacts monitored; community feedback incorporated

Table 2: Mapping challenges to framework components.

Challenge	Primary framework components	Secondary components
Algorithmic bias	Equity-Centered Validation (all components)	Anticipatory Governance (equity impact assessment, certification); Workforce Augmentation (AI literacy)
Privacy and surveillance	Community Sovereignty (data sovereignty, population-specific safeguards)	Anticipatory Governance (transparency, standards); Equity-Centered Validation (bias monitoring)
Reliability and hallucination	Equity-Centered Validation (continuous monitoring, local validation)	Anticipatory Governance (certification, transparency); Workforce Augmentation (human oversight)
Governance gaps	Anticipatory Governance (all components)	Community Sovereignty (engagement); Workforce Augmentation (capacity building)
Commercial interests	Anticipatory Governance (certification, standards)	Community Sovereignty (data sovereignty); Equity-Centered Validation (disaggregated evaluation)
Infrastructure constraints	Workforce Augmentation (capacity building)	Anticipatory Governance (tiered standards); Equity-Centered Validation (local adaptation)
Precaution-access trade-off	Anticipatory Governance (risk-tiered standards)	Equity-Centered Validation (phased validation); Workforce Augmentation (capacity building)

Workforce augmentation: EIOS includes training for member state users and emphasizes human analyst verification of AI-generated signals. Capacity building for LMIC users has been prioritized, though infrastructure constraints (connectivity, computing) limit access in some settings.

Community sovereignty: EIOS operates through national focal points, providing a degree of national ownership. However, community-level engagement is limited, and data sovereignty provisions are not systematically implemented.

Framework-informed recommendations: EIOS would benefit from: systematic equity impact assessments;

expanded disaggregated performance monitoring; enhanced transparency about algorithmic methods; strengthened community engagement mechanisms; and explicit attention to infrastructure requirements in LMIC deployment contexts.

Case 2: Preparedness Data Exchange (PDX)

PDX is an AI-enabled intelligence system for health security in Africa, synthesizing data across preparedness domains.

Anticipatory governance: PDX operates under WHO Africa governance with national ownership emphasized. The system includes an AI assistant with source-cited, auditable



outputs, supporting transparency. However, formal equity impact assessments and third-party certification are not documented.

Equity-centered validation: PDX was designed for African contexts with attention to local data sources and priorities. Validation has been conducted in multiple countries, with ongoing monitoring. Disaggregated performance evaluation is not systematically reported.

Workforce augmentation: PDX includes training for health officials and emphasizes augmentation rather than replacement of human judgment. Capacity building is integrated into implementation. Infrastructure constraints, including connectivity and electricity, are acknowledged in system design.

Community sovereignty: PDX emphasizes national ownership and capacity strengthening, reflecting data sovereignty principles. Community-level engagement is less developed, and population-specific safeguards are not systematically addressed.

Framework-informed recommendations: PDX would benefit from: formal equity impact assessments; systematic disaggregated performance reporting; expanded community engagement beyond national focal points; documentation of population-specific safeguards; and explicit protocols for addressing distribution shift.

Discussion

The Anticipatory Intelligence framework proposed here addresses a critical gap between the rapid proliferation of AI in public health and the governance, validation, and equity infrastructure necessary for responsible deployment. The framework synthesizes insights from global health research prioritization, practical implementations of AI surveillance systems, ethical analyses of bias and privacy, and technical literature on model validation to propose an integrated approach with explicit attention to resource-limited settings.

Several themes emerge across the framework's pillars. At the outset, AI in public health must be understood as a sociotechnical system, not merely a technical artifact. Technical performance metrics are necessary but insufficient indicators of responsible deployment. The structural context, including historical inequities, institutional capacity, community trust, and regulatory environment, determines whether technically capable systems produce equitable and beneficial outcomes. This is particularly salient in resource-limited settings, where infrastructure constraints (connectivity, electricity, computing, maintenance) can render technically capable systems non-functional or unsustainable.

Closely intertwined with this, the relationship between AI and human judgment is central. The framework consistently emphasizes augmentation rather than replacement,

recognizing that public health decisions involve values, contextual knowledge, and accountability relationships that cannot be delegated to algorithms. The WHO Africa formulation, that AI augments rather than replaces epidemiologists and public health leaders, captures this principle precisely. Operationalizing augmentation requires attention to workflow design, professional education, and organizational culture, as well as infrastructure that enables human oversight.

Running through all of this, equity must be an explicit design objective, not an afterthought. The framework embeds equity concerns throughout: in governance requirements for equity impact assessments and certification; in validation requirements for disaggregated performance reporting and local adaptation; in workforce requirements for bias literacy; and in community sovereignty requirements for data governance and population-specific safeguards. This integration reflects the recognition that AI systems optimized for average performance may systematically disadvantage marginalized populations, and that achieving equity requires intentional design rather than assuming it will emerge from technical excellence.

Beyond these considerations, governance must be anticipatory rather than reactive. The current pattern, deploy first, identify harms later, attempt remediation, imposes unacceptable risks on vulnerable populations and undermines trust in public health institutions. The framework's emphasis on pre-deployment equity assessments, third-party certification, and transparency requirements aims to shift the burden of proof from communities to demonstrate harm to developers and deployers to demonstrate safety and equity. However, this must be balanced against the imperative for timely access to beneficial technologies in resource-limited settings, requiring risk-tiered approaches and international collaboration.

Underpinning all of these themes, infrastructure and capacity constraints must be systematically addressed for AI to benefit resource-limited health systems. Connectivity, electricity, computing, interoperability, laboratory and registry capacity, device procurement and maintenance, and regulatory capacity are not merely implementation details but fundamental determinants of whether AI can be deployed equitably and sustainably. The framework addresses these constraints through: capacity building in LMIC contexts; design for offline operation and low-bandwidth connectivity; tiered standards that account for regulatory capacity; and phased implementation approaches.

The framework's limitations should be acknowledged. To begin with, it is a conceptual proposal requiring empirical validation through implementation and evaluation. The challenges of operationalizing equity impact assessments, establishing certification bodies, and building community



governance structures are substantial and context-dependent. A further limitation is that the framework addresses AI deployment within existing public health systems, but more fundamental questions about the direction of health system development, including the role of AI in market-driven versus public health-oriented models, remain beyond its scope. Compounding these concerns, the framework's recommendations, while grounded in evidence and ethical analysis, involve trade-offs that will require local deliberation. The appropriate balance between transparency and privacy, between standardization and flexibility, and between technical innovation and precaution will vary across contexts. Last but not least, the evidence base for AI in resource-limited settings remains limited, with most studies conducted in high-income contexts; this limits the generalizability of findings and underscores the need for research in LMIC settings.

Conclusion

Artificial intelligence holds genuine promise for strengthening public health capacity, enhancing surveillance, extending diagnostic expertise, optimizing resource allocation, and enabling anticipatory rather than reactive responses to health threats. The systems and applications reviewed in this paper demonstrate that this promise is not merely hypothetical. From the WHO's global open-source intelligence system to AI-enhanced decision support for elderly risk stratification, real-world implementations are already generating value. However, the promise of AI is contingent on addressing the infrastructure, capacity, governance, and equity challenges that characterize resource-limited health systems.

Yet the same capabilities that make AI promising also create risks. Algorithmic bias can exacerbate the health inequities that public health is committed to eliminating. Commercial data sources and surveillance capabilities can violate privacy and community sovereignty. Unreliable systems can produce misinformation and erode trust. And the prioritization of commercial interests over public health values can lead to deployment of tools that serve institutional efficiency rather than population health.

The Anticipatory Intelligence framework proposed here offers a path forward, not a set of technical fixes but a comprehensive approach to governance, validation, workforce development, and community engagement, with explicit attention to the infrastructure and capacity constraints characteristic of resource-limited settings. The framework's four pillars are interconnected: governance without validation is empty, validation without community sovereignty is extractive, workforce augmentation without governance lacks authority, and community engagement without validation is powerless. Only by addressing all four pillars together, and by investing in the infrastructure that enables their implementation, can public health institutions harness AI's potential while safeguarding the values of equity, transparency, and trust that underpin effective practice.

The stakes are high. Public health decisions shape life expectancy, health equity, and community resilience. AI will increasingly inform those decisions, whether through formal decision-support systems, automated surveillance, or the subtle influence of algorithmic prioritization. The question is not whether AI will shape public health practice; it already does, but whether that shaping will occur through deliberate, accountable, equity-centered design or through the haphazard adoption of tools optimized for other values.

The framework presented here represents a contribution to the deliberate path. Its implementation will require sustained commitment from policymakers, practitioners, developers, and communities. It will require resources for governance infrastructure, validation research, workforce development, and community engagement, as well as for the connectivity, electricity, computing, and maintenance infrastructure that AI systems require. It will require patience, as the framework's emphasis on anticipatory governance and local validation may slow deployment relative to the current pace of adoption. And it will require humility, as we acknowledge that even well-designed systems can fail and that ongoing learning and adaptation are essential.

Public health has a long history of balancing technological innovation with precaution, professional judgment with community engagement, and efficiency with equity. The integration of AI into public health practice requires the same balance. The Anticipatory Intelligence framework offers a structure for achieving it.

References

1. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. *BMJ*. 2020;369:m1328. Available from: <https://doi.org/10.1136/bmj.m1328>
2. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO; 2021.
3. World Health Organization. Epidemic Intelligence from Open Sources (EIOS): Annual report 2024. Geneva: WHO; 2024.
4. World Health Organization. AI for health: Accelerating progress towards universal health coverage. Geneva: WHO; 2024.
5. WHO Regional Office for Africa. From data to decisions: Strengthening health security in Africa. Brazzaville: WHO AFRO; 2024.
6. Linge JP, Steinberger R, Weber TP, Yangarber R, van der Goot E, Al Khudhairy DH, et al. Internet surveillance systems for early alerting of health threats. *Eurosurveillance*. 2009;14(15):19190.
7. Burrell J. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*. 2016;3(1):1-12.
8. Watts N, Amann M, Arnell N, Ayeb-Karlsson S, Beagley J, Belesova K, et al. The 2020 report of the Lancet Countdown on health and climate change. *Lancet*. 2021;397(10269):129-170. Available from: [https://doi.org/10.1016/s0140-6736\(20\)32290-x](https://doi.org/10.1016/s0140-6736(20)32290-x)
9. Bloom DE, Chen S, Kuhn M, McGovern ME, Oxley L, Prettnner K. The economic burden of chronic diseases in aging populations. *J Econ Ageing*. 2024;28:100512. Available from: <https://doi.org/10.1016/j.jeoa.2018.09.002>
10. WHO Regional Office for Africa. PDX data integration framework. Brazzaville: WHO AFRO; 2024.
11. Adirim T, Molten M. Artificial intelligence in public health: The need for stronger governance. *Am J Public Health*. 2024;114(4):601-605.



12. Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, Zittrain J, et al. The clinician and dataset shift in artificial intelligence. *N Engl J Med*. 2021;385(3):283–286. Available from: <https://doi.org/10.1056/nejmc2104626>
13. Cedars-Sinai. Artificial intelligence advances in cardiovascular medicine. Los Angeles: Cedars-Sinai Medical Center; 2024.
14. World Health Organization. Global health security: Challenges and opportunities for the post-COVID era. Geneva: WHO Press; 2024.
15. Belizaire MR. Interview on AI-enabled preparedness. WHO Africa Newsletter. March 2024.
16. The Lancet Global Health. Research priorities for data science and artificial intelligence in global health: An international consensus exercise. *Lancet Glob Health*. 2024;12(3):e401–e412.
17. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: Implications for health systems. *J Glob Health*. 2019;9(2):020318. Available from: <https://doi.org/10.7189/jogh.09.020318>
18. WHO Regional Office for Africa. Preparedness Data Exchange: Technical documentation. Brazzaville: WHO AFRO; 2024.
19. Wachter S, Mittelstadt B, Floridi L. Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *Int Data Priv Law*. 2017;7(2):76–99.
20. WHO Regional Office for Africa. From data to decisions: Strengthening health security in Africa. Brazzaville: WHO AFRO; 2024.
21. WHO Regional Office for Africa. From data to decisions: Strengthening health security in Africa. Brazzaville: WHO AFRO; 2024.
22. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56.
23. Char DS, Abràmoff MD, Feudtner C. Identifying ethical considerations for machine learning healthcare applications. *Am J Bioeth*. 2020;20(11):7–17. Available from: <https://doi.org/10.1080/15265161.2020.1819469>
24. World Health Organization. EIOS 2.0: Enhanced epidemic intelligence for global health security. Geneva: WHO; 2024.
25. Cedars-Sinai. Molecular Twin: Personalized cancer treatment through AI. Los Angeles: Cedars-Sinai Medical Center; 2024.
26. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: How can AI contribute to health in resource-poor settings? *BMJ Glob Health*. 2018;3(4):e000798. Available from: <https://doi.org/10.1136/bmjgh-2018-000798>
27. Chen X, Liu Y, Wang H, et al. Advancing personalized healthcare for elderly individuals through AI-enhanced preventive disease risk stratification. *Decis Support Syst*. 2024;185:114201.
28. WHO Regional Office for Africa. AI assistant for preparedness intelligence: User guide. Brazzaville: WHO AFRO; 2024.
29. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94–102. Available from: <https://doi.org/10.7861/futurehosp.6-2-94>
30. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med*. 2018;178(11):1544–1547. Available from: <https://doi.org/10.1001/jamainternmed.2018.3763>
31. Eubanks V. Automating inequality: How high-tech tools profile, police, and punish the poor. New York: St. Martin's Press; 2018.
32. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. 2019;25(1):70–74. Available from: <https://doi.org/10.1038/s41591-018-0240-2>
33. Hay SI, George DB, Moyes CL, Brownstein JS. Big data opportunities for global infectious disease surveillance. *PLoS Med*. 2013;10(4):e1001413. Available from: <https://doi.org/10.1371/journal.pmed.1001413>
34. The Lancet Global Health. Research priorities for data science and artificial intelligence in global health: An international consensus exercise. *Lancet Glob Health*. 2024;12(3):e401–e412.
35. Tiller D, Whyte D, Bickmore T. Assessing the accuracy of generative AI chatbots in health-related queries. *J Med Internet Res*. 2024;26(1):e45678.
36. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402–2410. Available from: <https://doi.org/10.1001/jama.2016.17216>
37. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019. *Lancet*. 2020;396(10258):1223–1249. Available from: [https://doi.org/10.1016/s0140-6736\(20\)30752-2](https://doi.org/10.1016/s0140-6736(20)30752-2)
38. Rivers C, Chretien JP, Riley S, et al. Using “outbreak science” to strengthen the use of models during epidemics. *Nat Commun*. 2019;10:3102.
39. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—Addressing ethical challenges. *N Engl J Med*. 2018;378(11):981–983. Available from: <https://doi.org/10.1056/nejmp1714229>
40. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*. 2019;1(5):206–215. Available from: <https://doi.org/10.1038/s42256-019-0048-x>
41. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115–118. Available from: <https://doi.org/10.1038/nature21056>
42. Sims N, Kasprzyk-Hordern B. Future perspectives of wastewater-based epidemiology. *Water Res*. 2020;171:115456.
43. WHO Regional Office for Africa. Preparedness Data Exchange: Technical documentation. Brazzaville: WHO AFRO; 2024.
44. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*. 2019;17(1):195. Available from: <https://doi.org/10.1186/s12916-019-1426-2>
45. Chen PT, Wu T, Wang P, et al. Pancreatic cancer detection from CT scans using deep learning. *Radiology*. 2024;310(2):e231456.
46. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447–453. Available from: <https://doi.org/10.1126/science.aax2342>
47. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med*. 2018;1:18. Available from: <https://doi.org/10.1038/s41746-018-0029-1>
48. Adirim T, Molten M. Artificial intelligence in public health: The need for stronger governance. *Am J Public Health*. 2024;114(4):601–605.
49. World Health Organization. Epidemic Intelligence from Open Sources: User statistics and impact assessment. Geneva: WHO; 2024.
50. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: Towards better research applications and clinical care. *Nat Rev Genet*. 2012;13(6):395–405. Available from: <https://doi.org/10.1038/nrg3208>
51. The Lancet Global Health. Research priorities for data science and artificial intelligence in global health: An international consensus exercise. *Lancet Glob Health*. 2024;12(3):e401–e412.
52. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: How can AI contribute to health in resource-poor settings? *BMJ Glob Health*. 2018;3(4):e000798.
53. International Telecommunication Union. Measuring digital development: Facts and figures 2024. Geneva: ITU; 2024.
54. World Bank. Tracking SDG 7: The energy progress report 2024. Washington, DC: World Bank; 2024.