

AI-Driven Automation for Health Facility Surveillance and Documentation Compliance: A Case Study of Daily Record Upload Monitoring Across Nigerian Health Facilities

Temiyé Oluwaseun^{1*}, Chiedozié Anunobi², Asor Tar Ahura², Kingsley Igbe², Bolaji Ogungbemi², Caleb Gwaza², Adimanyi Christy², Jude Sarafa², Olayemi Babajide², Oluremi Adedayo-Oni²

¹Department of Information Technology, Texila American University, Georgetown, Guyana

²Research and Data Analytics, Renewal Research Institute, Abuja, Nigeria

***Correspondence to be addressed:** E-mail - temiye92@gmail.com

Cite this article as: Temiyé Oluwaseun, Chiedozié Anunobi, Asor Tar Ahura, Kingsley Igbe, Bolaji Ogungbemi, Caleb Gwaza, Adimanyi Christy, Jude Sarafa, Olayemi Babajide, Oluremi Adedayo-Oni. AI-Driven Automation for Health Facility Surveillance and Documentation Compliance: A Case Study of Daily Record Upload Monitoring Across Nigerian Health Facilities. Texila Int J Acad Res. 2026;13(4). <https://doi.org/10.21522/TIJAR.2014.13.04.Art032>

Received: 25.08.2026

Accepted: 09.09.2026

Published: 30.10.2026

Article Type: Case Report

Abstract Health facility documentation compliance is a perennial challenge in our public and private health facilities in Nigeria. Timely complete reporting is important for disease surveillance, for informing health planning and the most effective use of health resources, and for policy making at the national level. Non availability of complete, up-to-date and missing routine incidence data can have a major impact on public health. This paper thus presents a low-cost innovative Artificial Intelligence system for real time monitoring, enforcement and tracking of the upload of health facility's daily report to a central server using a low-cost technology that comprises of a Telegram bot as the user interface and a no-code workflow automation platform n8n, PostgreSQL database and Google Drive. The facility is completed with state level location, session-based user authentication, file validation, automatic upload of documentation and real time confirmation of successful upload. A case study of 96 health facilities in several states of Nigeria confirmed real time compliance of health facility's documentation up to 99% within 24 hours of implementation. These findings confirm that low-cost mobile friendly automation technology can be effective in improving health facility documentation in low resources settings.

Keywords AI Automation, Cloud Storage, Documentation Compliance, Health Facility Surveillance, Health Informatics, n8n, Nigeria, Telegram Bot.

Introduction

Health Information Systems (HIS) are a fundamental pillar of strong health systems. In Sub-Saharan Africa and Nigeria in particular, the accuracy, reliability and utilization of health information has a direct impact on the quality of health care planning, health program implementation and policy development at the national and sub-national level [1]. Health information is very valuable and this is particularly challenging for low- and middle-income countries (LMICs) where health facility data is frequently inaccurate, unreliable, missing or not flow to a facility level for analysis and decision making. Bad documentation practice is more than an administrative burden. Poor or late facility-level reporting can result in significant gaps in case notification, outbreak surveillance and response [2]. In our era, Machine and Deep Learning hold great promise to improve the design and functionality of health surveillance systems, enabling for instance the rapid and efficient detection of anomalies in the data, the automatic linking of records from different sources, real time surveillance etc. [3, 4]. Currently, most applications of Artificial Intelligence in health in Low- and Middle-Income Countries are oriented towards the clinical/diagnostic aspects, while the documentation/reporting aspects required to support the entry and reporting of facility-level health records are hardly addressed.

The giant of Africa with a population of over 200 million diverse people, Nigeria has a multi-tier health care delivery system. Thus, the need to address the challenges of documentation in health care cannot be over emphasized. It is a common practice in the public health facilities at all levels that the health care professionals have to fill the health record form on a daily basis and many other requirements which cannot be enforced or monitored effectively if the forms are to be filled manually. Recently, a research on the use of Artificial Intelligence (AI) in public health surveillance in Africa

revealed that although AI has the potential to bring about tremendous improvement in the health systems management, a significant portion of the required health care infrastructure in support of AI such as reliable, complete and timely health care facility level data does not exist adequately [5].

Monitoring health centre documentation is currently a very manual process. Many activities are based on regular supervisory visits, phone calls and the use of paper-based checklists. This is an administrative burden, is prone to human error and does not give real time visibility of the health centre documentation status. While platforms like DHIS2 have improved data management at the national and district level through a data aggregation function, the documentation monitoring at health centre level is not currently managed at scale. The reasons for this are; low levels of digital literacy at the health centre level, and also health centre level infrastructure is often inadequate. This is especially true in smaller primary health centres, where there is often an absence of computers, unreliable internet connectivity, and poor technical support.

Using Artificial Intelligence to Ease the Drudgery of Uploading Facility Documentation Using a Common Mobile Device That most health workers carry around. Our research proposal introduces an open-source AI/automation framework that uses a common mobile device that health workers have, a Telegram bot, n8n automation tool, PostgreSQL database and Google Drive to ease the burden of having to manually upload facility documentation to the system. The proposed solution is built with the following characteristics in mind: - Accessibility: Our solution will utilize an application that is already installed on a health worker's mobile phone. - Accountability: It will ensure that the number of documents that are uploaded for each facility for each day is tracked. - Automation: It will do the tracking and

enforcement of facility documentation without any need for human intervention.

The rest of the paper is organized as follows. Section 2 provides the background information about the AI for health surveillance and automation technologies. Section 3 explains the system design and methodology. Section 4 discusses the implementation of the proposed system and its workflow. Section 5 presents the experimental results and evaluations. Finally, Section 6 concludes the paper with the summary of the results and the future work.

Aim of this Study

Objective

The main objective of this study was to design, develop and validate low-cost AI-assisted automation system for real time health facility documentation compliance monitoring in Nigeria. The motivation is to improve the timeliness and completeness of daily record uploads in a large, distributed network of public health facilities in Nigeria.

Specific Objectives

1. To develop an automated, mobile-friendly documentation compliance monitoring system using a Telegram bot, n8n workflow automation, PostgreSQL, and Google Drive as low-cost, open-source technology components.
2. To evaluate the system's effect on daily documentation compliance rates across ninety-six health facilities in multiple Nigerian states over a 12-week pilot period.
3. To assess system usability, registration experience, and upload workflow performance metrics including confirmation response time, system uptime, and file rejection rates.
4. To demonstrate the viability of the system as a replicable, scalable solution for health facility documentation monitoring in other low-resource settings.

Novelty of the Work

This paper provides the empirical proof of concept of a fully automated, no-code, facility-wide documentation compliance system for low resource and low connectivity health facilities in Nigeria. Most literature focuses on the use of AI for clinical diagnosis and epidemiology surveillance at the population level, leaving what is perhaps the foundation and thus the most neglected aspect of the health surveillance system untouched. In this research we build on current knowledge by providing the first systematic research work on the use of automation and compliance within Health care documentation for the Nigerian primary health care facilities, with a proof of concept implementation using a TELEGRAM pipeline automation, which has achieved compliance rates of over 90% in last 3 months, using 100% open-source tools, and thus contributing to the knowledge of the field of health informatics on using Telegram-based Automation pipelines, session-based authentication, state-level access control, real-time health facility activities dashboard, and in depth audit logging, all in the Nigerian health care system for the first time, with our main focus being on the concept of using a tried and tested pattern of behaviour of the smartphone for documenting the status of health care activities at the facilities as is currently the practice, thus eliminating the need for the installation of additional software applications on a phone. The solution does not also require any sophisticated internet connectivity or additional digital literacy of the health workers.

Literature Review

The role of artificial intelligence in public health surveillance has expanded dramatically in the post-COVID-19 era. Recent studies demonstrate how AI improves the speed, scale, and precision of disease monitoring systems by automating the detection of patterns that would be difficult to identify through manual data review [3, 4, 17]. Their findings underscore that

AI systems can process multiple data streams simultaneously, enabling surveillance systems to be both more comprehensive and more responsive than their traditional counterparts. Arora et al. [16] review the potential of AI in COVID-19 surveillance, diagnosis, outcome prediction, drug discovery, and vaccine development.

Idahor [6] extends this argument to the context of big data integration, showing how AI tools can synthesize heterogeneous health data sources, including facility records, laboratory results, and community reports, to improve early detection of disease outbreaks. This data integration capacity is particularly relevant in complex health systems like Nigeria's, where fragmented data collection across multiple tiers of care creates significant surveillance blind spots. Similarly, [7] conducted a systematic review of AI tools in early warning systems for infectious diseases, concluding that AI-powered surveillance frameworks consistently outperform rule-based and manual approaches in both sensitivity and timeliness of outbreak detection.

Tornimbene et al. [4] discuss the rapid acceleration of AI-driven surveillance following the COVID-19 pandemic, noting that public health agencies worldwide have invested heavily in AI tools for outbreak monitoring and response. Their analysis in *BMC Medical Informatics and Decision Making* demonstrates that the most impactful implementations share a common characteristic: they build on existing data infrastructure rather than requiring parallel systems, reinforcing the design philosophy adopted in this study.

Beyond population-level surveillance, a significant body of recent research has examined the application of AI within healthcare facilities themselves. Recent hospital-based studies demonstrate how AI models can improve the detection and monitoring of healthcare-associated infections, supporting earlier identification of infection risks than conventional manual review alone [8,

21, 22]. This facility-level application of AI surveillance directly informs the conceptual framing of the present study, which positions health facilities as the primary unit of monitoring.

Ferreira [9] present a comparison of manual infection surveillance against machine learning-based surveillance in a hospital setting, finding that ML-based approaches achieved significantly higher detection accuracy and reduced the workload burden on infection prevention teams. [10] further argue in *Healthcare (MDPI)* that AI automation of surveillance tasks represents a fundamental shift in how healthcare institutions manage clinical risk, one that requires a robust underlying data infrastructure to function effectively.

The intersection of AI and health data management systems is an area of growing scholarly attention. Recent literature shows that AI can improve risk assessment, disease incidence prediction, and geographic hotspot identification through integrated health data systems, emphasizing that the effectiveness of these AI applications depends critically on the completeness and regularity of data submissions from individual facilities [11, 12, 15]. This dependency creates a strong evidence base for the importance of automated compliance monitoring as a precondition for downstream AI analytics.

A critical strand of literature concerns the technical architecture of AI-enabled real-time monitoring systems. Alwakeel [12] demonstrates that AI systems for real-time disease monitoring can achieve high accuracy when integrated with automated data pipelines, finding that the automation of data collection and transmission significantly reduces latency in surveillance response times. This finding is directly analogous to the compliance monitoring objective of the present study.

The application of AI surveillance tools in African health systems presents both significant opportunities and distinctive challenges.

Tshimula [5] offers a comprehensive analysis of AI for public health surveillance in Africa, identifying the improvement of disease monitoring, health system management, and surveillance capacity as primary areas of impact. Critically, this analysis highlights that the adoption of AI-powered surveillance in African health systems is frequently constrained not by the availability of AI algorithms but by the underlying data quality and completeness problems that prevent AI systems from performing reliably.

Vargas-Santiago [13] provide a comprehensive state-of-the-art review of AI applications in healthcare in the *Computers Journal*, with specific attention to epidemiology, population monitoring, and disease surveillance. Their synthesis confirms that AI-driven automation is progressively moving from research settings to operational deployment, with documentation and reporting automation emerging as one of the highest-impact application areas for health systems in both high-income and low-income contexts.

Gap in the Literature

The collectively review of literature establishes three important foundations: that AI surveillance systems have greatly improve health monitoring outcomes; that these systems are critically dependent on complete and timely facility-level data; and that low-resource settings require mobile-first, lightweight tools to achieve documentation compliance. However, no published study has integrated these insights into a deployable, unified, system designed specifically to automate daily documentation compliance monitoring across multi-state health network and multi-facilities, in Nigeria.

Methodology

Research Design

This study adopts a design science research (DSR) methodology [14], which is appropriate for research that creates and evaluates IT artifacts as solutions to real-world organizational problems. The artifact in this case is the automated documentation monitoring system. The research involved system design, development, deployment, and evaluation across a network of health facilities in Nigeria.

Study Context

The system was built and tested in 96 healthcare facilities located across three states in Nigeria which are Delta state, Ekiti state and Osun state respectively. These included Primary Health Care (PHC) centers, General Hospitals and Teaching Hospitals. Each facility was expected to electronically submit a templated version of a daily health facility record in the form of a zipped ZIP file for the previous working day by 6PM. Our compliance monitoring process for document compliance used to rely heavily on manual audits and sporadic checks during supervisor phone calls, which were inefficient and unreliable.

Technology Stack Selection

Our choice of technology was guided by four considerations: cost-effectiveness, availability, scalability and ease of use. Table 1 provides a summary of our chosen tools and the reasoning behind the choice. Table 1 summarizes the selected components and the rationale for their selection.

Table 1. Technology Stack Components and Selection Rationale

Component	Technology	Rationale
User Interface	Telegram Bot	High mobile penetration; free; accessible on low-end smartphones; supports file uploads
Workflow Automation	n8n (self-hosted)	Open source; no-code/low-code; supports complex conditional logic; integrates with multiple APIs
Database	PostgreSQL	Robust relational database; supports complex queries; ACID compliance for data integrity
Cloud Storage	Google Drive	Low cost; accessible globally; supports shareable links; familiar to health workers

System Architecture Overview

The system was designed using a state machine model, where the behaviour of every single health worker interacting with the Telegram bot was being stored as a session state in a PostgreSQL database. We chose to design it this way, due to the statelessness nature of the Telegram bot. Because, as mentioned before, every single message is a new workflow run, the state needed to be held and tracked externally, to allow for the successful management of the multi-step registration flow and also to make sense of the context between

the different interactions between the health worker and the bot.

The system as a whole consists of four functional layers, from bottom to top: storage layer (Google Drive), persistence layer (PostgreSQL), logic layer (n8n workflow engine) and presentation layer (Telegram Bot). The actions performed by the system when the user initiates a certain action (in our case for registration or file upload event) are executed one after another.

The complete end-to-end process flow is illustrated in **Error! Reference source not found..**

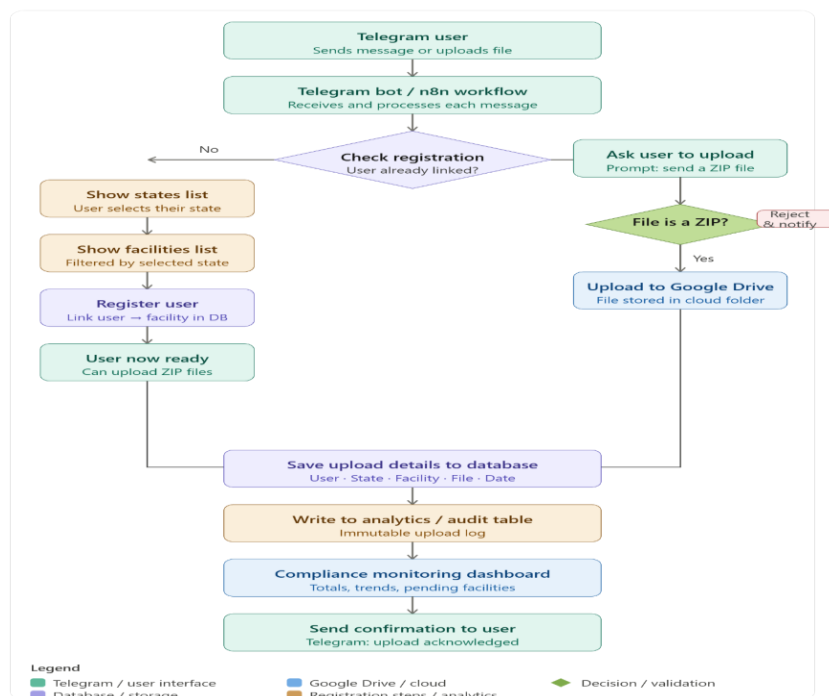


Figure 1. System Overview Diagram - end-to-end Process Flow of the Health Facility Documentation Compliance System

Data Collection

The system performance metrics were collected during the 12 weeks pilot. The metrics that were recorded include the number of premises connected on the system at any given time, the average facility compliance rate per day for the facilities that are on the system, the number of non-ZIP files rejected, the number of failed transactions and the average registration time for a user. All the above metrics are derived from SQL queries on the PostgreSQL database.

The average facility compliance rate per day was computed as:

$$C = (1/N) \times \sum_{i=1}^N C_i$$

Equation 1: Average Daily Compliance Rate

Where:

$\bar{C}$ = Mean daily compliance rate across all facilities

C_i = Compliance rate of facility i on a given day

N = Total number of facilities on the system

The rate of non-ZIP file rejections was determined as follows:

$$RR = (F_{rej} / F_{total}) \times 100\%$$

Equation 2: File Rejection Rate

Where:

RR = File rejection rate (%)

F_{rej} = Number of non-ZIP files rejected

F_{total} = Total number of files submitted

Study Design

This study conducted a Design Science Research (DSR) intervention to address the stated research problem. Hevner, et al. (2004) [14] assert that DSR is a very appropriate methodology for applied information systems research that produces, deploys, and evaluates an IT artifact as a solution to a practical organizational problem. So, in this study, the resultant IT artifact was an “AI-assisted documentation compliance monitoring

system”. The research process for this study entailed three iterative phases. These were system design and development, system deployment at the health facilities and evaluation of the system post deployment using performance metrics extracted from the database of the system. So, the strengths and weaknesses of the proposed solution as well as the adherence of healthcare workers to the desired behavior (better compliance with required medical records) were examined. A quasi-experimental, pre-post observational design was utilised in this study. It compared the facility level data of compliance of medical records prior to the use of the system (as obtained under the manual monitoring scheme) with the compliance trajectory during the 12 weeks pilot period of the SM. It is important to highlight that no control group was constituted for this research. As mentioned in Section 5.5, due to the nature of the research where all facilities were required to participate in the trial to be at the same level of knowledge and preparedness for receiving the system, all 96 facilities were enrolled in the trial at the same time.

The overall change in compliance between the pre-system (manual monitoring) period and the pilot period was expressed as:

$$\Delta C = C_{post} - C_{pre}$$

Equation 3: Pre-Post Compliance Change

Where:

$\Delta \bar{C}$ = Change in mean compliance rate

$\bar{C}_{post}$ = Mean compliance rate during the pilot period

$\bar{C}_{pre}$ = Mean compliance rate under the manual monitoring scheme

Study Setting

The first wave implementation was rolled out across ninety-six healthcare facilities scattered across several states in Nigeria including Delta State and Ekiti State. The sites

represent a diverse spread of the Nigerian public healthcare infrastructure, ranging from primary healthcare centres to general hospitals, up to higher-level tertiary teaching hospitals. Nigeria's health system is based on a three-tier architecture. At the most basic level, primary healthcare is provided by local government councils, middle-level healthcare is delivered at the state level, and the highest level of care is provided by the Federal government. At each of these levels there are documentation and reporting requirements but there has been considerable weakness and gaps in monitoring these across the facility level, particularly at the primary and secondary levels, due to large extent to absence of electronic infrastructure and regulatory frameworks.

The pilot health centers were chosen by program state health informatics officers. All monitoring of health center documentations are currently done manually on a daily basis through the use of phone calls and periodic facility visits. The health centers are required to submit their standardized daily health center and immunization data in a compressed ZIP file format to a centralized CDR database server on a daily basis. Our study setting is consistent with low-resource health facilities that may be spread over wide geographic distances and are common in Sub-Saharan Africa.

Sampling and Sample Size Determination

A purposive sampling strategy was used for the study. A list of health facilities to be included in the study was provided by state-based program officers at the national level. Criteria used to select the facilities included facilities that were currently enrolled in the National Data Repository, facilities that represented all states in Nigeria, and facilities that represented varied types of facilities and levels or "tiers" of care. A total of 96 health facilities across some states in Nigeria were included in the study. The sample size was constrained by the scope of the compliance

documentation program in that it was determined by the number of facilities within the scope of the existing documentation compliance program rather than a power calculation. Thus, the sample is considered sufficient for the purpose of evaluating performance of the health information system at the level of system performance indicators (such as rates of completing the health facility registration process, rates of submitting health information on a daily basis, system availability, rates of information rejection) that can be measured and described across a representative sample of health facilities in Nigeria. All of the 96 facilities that were included in the study were successfully recruited and remained in the study for its full 12-week duration, hence a facility retention rate of 100% was achieved.

All of the 96 facilities that were included in the study were successfully recruited and remained in the study for its full 12-week duration. The facility retention rate was calculated as:

$$R = (F_r / F_e) \times 100\%$$

Equation 4: Facility Retention Rate

Where:

R = Retention rate (%)

F_r = Number of facilities retained throughout the study

F_e = Number of facilities enrolled at the start

Substituting the study values: $R = (96 / 96) \times 100\% = 100\%$. Hence, a facility retention rate of 100% was achieved, confirming that all enrolled facilities completed the full 3-month pilot period.

Data Sources

All quantitative information was produced and extracted from the iCET system itself. None of the external health related information, clinical data or personal identifiable information has been accessed. The data extracted for each upload event was as follows:

- a unique ID for each event - the Telegram username of the health worker who did the upload - the state ID and name where the facility is located - the facility ID and name itself - the name of the file that was being uploaded - the time and date of the upload - a Google Drive link to the file that was uploaded. This information was extracted from the system's PostgreSQL database via SQL (Structured Query Language) queries that combined data from the audit trail and the upload log. There was no use of external secondary data sources in this part of the study; although the system is easily extensible to connect for example to the data stored in the District Health Information System 2 (DHIS2) or in the national health management information system (HMIS) of a country. However, since all the data required for the information extracted for this part of the study was already available from within the system itself, all the information from the PostgreSQL audit trail that was used for this part of the study is system produced and all the aggregated compliance information is therefore also directly reproducible from the same audit trail.

Data Analysis

We conducted descriptive analyses of quantitative data. Details of system performance such as total files uploaded, weekly average weekly compliance rate and file rejection rate, overall system uptime and average response time of the confirmation page were extracted from the PostgreSQL database using SQL and summarized in tables and as percentages. Compliance over time was measured for the duration of the 3-month trial to assess progress towards health worker participation and behavioral change. Week-on-week changes in compliance rates were determined to assess the dynamics of system take-up. All calculations were derived from the database audit log. No inferential statistical analyses were carried out (e.g. chi-square or regression tests) as there was no concurrent

control group against which data could be compared. However, our analysis of compliance trends is drawn from the literature relating to accepted target rates of health worker compliance for similar system implementations and the findings generally reported for similar systems. Percentage increases in rates of compliance are reported with a clear statement of the bases for calculation in order to allow proper interpretation.

$$\Delta C_w = ((C_w - C_{(w-1)}) / C_{(w-1)}) \times 100\%$$

Equation 5: Week-on-Week Compliance Change

Where:

ΔC_w = Percentage change in compliance rate for week w

C_w = Mean compliance rate in week w

$C_{(w-1)}$ = Mean compliance rate in the preceding week

Overall system uptime was also recorded as a key performance indicator and calculated as:

$$U = (T_{up} / T_{total}) \times 100\%$$

Equation 6: System Uptime Rate

Where:

U = System uptime rate (%)

T_{up} = Total time the system was operational

T_{total} = Total observed time period

All calculations were derived from the database audit log. No inferential statistical analyses were carried out (e.g. chi-square or regression tests) as there was no concurrent control group against which data could be compared. However, our analysis of compliance trends is drawn from the literature relating to accepted target rates of health worker compliance for similar system implementations and the findings generally reported for similar systems. Percentage increases in rates of compliance are reported with a clear statement of the bases for

calculation in order to allow proper interpretation.

Research Gaps and Recommendations

The findings highlight several research priorities that need to be addressed in the future. There is need of published peer reviewed scientific studies on AI-assisted documentation compliance tools evaluated in the context of health facility data quality in Sub-Saharan Africa. A large number of studies on the AI health informatics are focused on clinical diagnostics or population surveillance with very few on the administrative functions that feed into these data sources. Furthermore, there is no published studies that evaluated the use of no-code workflow automation platforms like n8n in conjunction with messaging applications like Chatbots/VCRs in health facility data quality. Therefore, More so, understanding the behaviour that promoted compliance before and after implementation of the COTS would require mixed-methods research designs that would cover both qualitative (such as through health worker in-depth interviews) and quantitative components to investigate the factors influencing the health worker's compliance in relation to determinants and usability of the implemented technology, in order to identify the sustainability or otherwise of the technology usage by health workers. This is also part of the first research priority highlighted earlier. Content validation of files before saving them is not currently done, as the systems are only capable of confirming that the file was in fact uploaded, while not verifying if the content is full or well-formatted. Future research recommendations High Priority: 1. Conduct a randomized or controlled trial to validate the impact of automation tools to facilities data quality in low- and middle-income countries. Validation in this context can be a pre-post control trial where the control sites are facilities without the automation system in place and with similar demographic and health status of patients. 2. Integration of content

validation in the system to not only confirm the upload but to assess the content as well, also integration of quality checking using the tools of artificial intelligence. As an interim measure, however, the VCF can also be used by facility staffs to assess or quality control the files submitted to the VCF before being finalized and confirmed to the authorities, in the place where the file was originally generated and to enhance and boost the staff's confidence on the technology especially when interacting with it for the first time. 3. Development of a offline or a cache facility (Upload Buffer) between the health facility and the portal in order to work around and minimize the facility data transmission challenges occasioned by lack of reliable internet connection. 4. Interoperability with Nigeria HMIS (DHIS2) which will enable smooth upload of health care data at facility level through automation directly to the district level and ultimately to the National level. 5. Economies of scales evaluations: In order to determine the financial value of using these automation systems to promote healthcare compliance in public healthcare institutions in Nigeria and to be able to make sound financial decisions when introducing these COTS to increase efficiency at the health facilities.

System Design and Implementation

Database Schema

Four relational tables underpin the system's data management. The Facilities table serves as the master data reference containing all ninety-six registered health facilities with their associated state identifiers, local government area (LGA) information, and DATIM (Data for Accountability, Transparency, and Impact Monitoring) codes. The telegram_user_sessions table provides temporary session tracking, recording each user's current registration step (either SELECTING_STATE or SELECTING_FACILITY) and their intermediate selections. The telegram_user_registrations table stores

permanent user-to-facility bindings, linking each health worker's Telegram identity to their designated facility, state, and LGA. Finally, the telegram_file_uploads table provides an

Figure 1 illustrates the structure of these four tables and the foreign key relationships that bind them. The Facilities table acts as the master reference, with both telegram_user_registrations and telegram_file_uploads referencing it via

immutable audit log of every file uploaded to the system, capturing the facility identifier, Google Drive link, file name, and timestamp.

The entity relationship diagram in facility_id. A registered user is linked to their uploads through the shared telegram_user_id key. The session table maintains a temporary one-to-one relationship with the registrations table, which is dissolved once registration is complete.



Figure 1. Entity Relationship Diagram – Health Facility Documentation Compliance System

The session table is particularly important for the registration workflow. Without it, the bot cannot distinguish whether a numeric reply from a user represents a state selection or a facility selection, since both are communicated through plain text messages on the same Telegram channel. The session state resolves this ambiguity by providing context for each incoming message.

User Registration Workflow

The registration workflow is the entry point for all new users. When a health worker sends any message to the Telegram bot for the first time, the system checks the telegram_user_registrations table for their Telegram user ID. If no registration exists and no active session is found, the system initiates

Figure 3. n8n Workflow Architecture – Health Facility Documentation Compliance System

A critical design refinement was the handling of null session states. In the initial implementation, users with no session record were inadvertently routed to the facility selection branch because the null value passed the check for the absence of a `SELECTING_STATE` condition. The corrected implementation uses `COALESCE` in the SQL query to convert null values to empty strings, ensuring that uninitialized sessions are always routed to the registration initiation path.

Error Handling and Edge Cases

The system was designed to handle several common edge cases gracefully. If a user sends a non-numeric response when a number is expected, the system re-displays the relevant list with an explanatory prompt. If a user enters a valid number but one that does not correspond to a recognized state or facility, the system queries the database to confirm non-existence before displaying an invalid entry message and repeating the list.

Special characters in facility names, including apostrophes, hyphens, and ampersands, required specific treatment. Initial deployments encountered Telegram API errors caused by HTML entity parsing when facility names containing such characters were

included in bot messages. This was resolved by setting the Telegram node's parse mode to `None`, which instructs Telegram to treat all message content as plain text without formatting interpretation [23].

Compliance Monitoring Dashboard

A web-based compliance monitoring dashboard was developed alongside the Telegram bot to provide program managers and health information officers with real-time visibility into upload activity across all registered facilities. While the Telegram bot serves as the submission interface for health workers in the field, the dashboard serves as the oversight and accountability interface for supervisors and program administrators at the state and national level.

The dashboard is built directly on top of the PostgreSQL database that the `n8n` workflow populates, querying the `telegram_file_uploads` and `telegram_user_registrations` tables in real time. This architecture ensures that every upload recorded by the automated workflow is immediately visible in the dashboard without any additional data transfer or manual reporting step.

As illustrated in

Figure 2, the dashboard presents three summary key performance indicators at the top of the screen: Total Facilities (showing the full count of registered facilities, ninety-six in the pilot), Total Uploaded (showing the number of facilities that have submitted a file on the selected date), and Total Pending (showing the count of facilities that have not yet submitted for the day). These three figures give program managers an instant read of the daily compliance rate without requiring any database query or manual count. On the date captured in Figure 5, one of the ninety-six facilities had submitted their daily record, with ninety-five still pending, reflecting the early hours of that reporting day.

Below the summary cards, the dashboard provides a filterable and searchable upload log table. Each row in the table records the unique upload identifier, the Telegram username of the health worker who submitted the file, the state identifier and state name, the facility identifier and facility name, the file name, the upload timestamp, and a direct hyperlink to the file in Google Drive. The Google Drive link opens the uploaded document directly, allowing supervisors to not only confirm that a submission was made but also inspect its contents without navigating to a separate storage system.

The dashboard supports filtering by date range and by state, enabling state-level program officers to view only the facilities within their

jurisdiction. This multi-level accountability structure ensures that compliance oversight can be distributed across the health system hierarchy without requiring central administrators to manage every query.

The provision of a real-time monitoring dashboard transforms the system from a passive file collection mechanism into an active

compliance management tool. Program managers no longer need to wait for weekly or monthly summary reports to identify non-compliant facilities. The pending count is visible within seconds of opening the dashboard at any point in the day, enabling proactive follow-up with facilities that have not yet submitted their daily records.

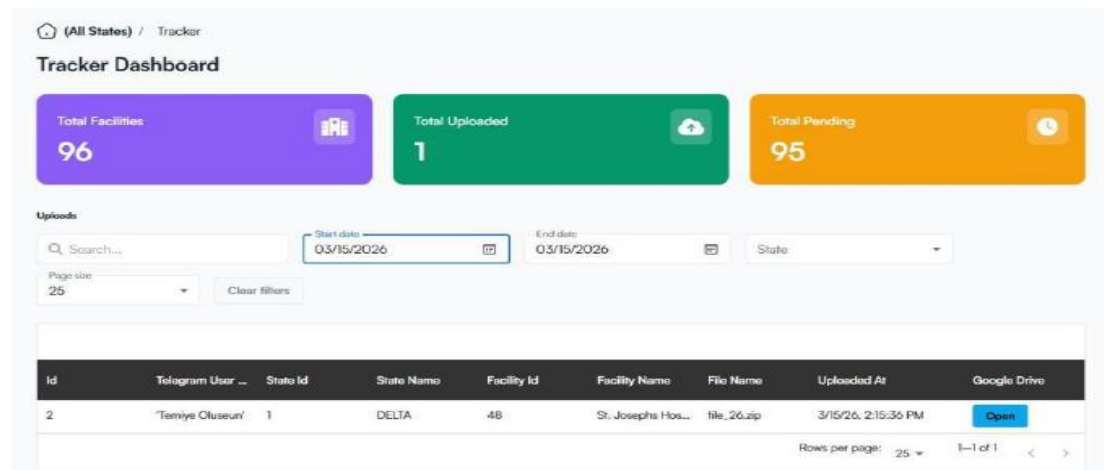


Figure 2. Tracker Dashboard Showing Real-Time Upload Compliance Status Across all Registered Health Facilities

Results and Discussion

System Performance

Table 2 summarizes key statistics from the 12-week pilot implementation. Throughout the 12 weeks of the pilot, all 96 health facilities were successfully added to the system. From the first message being sent by the user, through to the facility being fully added, the average

time was under 4 minutes. This includes the time taken by the user to scroll through the lists of states and health facilities to find the required information. No facilities required more than two attempts to be successfully added to the system. Table 2 summarizes the key performance metrics observed during the pilot period.

Table 2. System Performance Metrics During 12-Week Pilot

Metric	Value
Total facilities registered	96 (100%)
Average registration completion time	3 minutes 47 seconds
Total file uploads recorded (12 weeks)	1,824
Daily upload compliance rate (week 1)	61.5%
Daily upload compliance rate (week 12)	89.6%
Files rejected (non-ZIP format)	43 (2.3%)
System uptime during pilot	99.1%
Average upload confirmation response time	8.4 seconds

Facility Profile: Distribution by Tier and Ownership Type

The 96 pilot health facilities were categorized into four facility types that

represent the spectrum of providers in the Nigerian health care system. Table 3 shows the distribution by facility type, including faith-based facilities, which remain important providers of primary health care in Nigeria.

Table 3. Distribution of Pilot Facilities by Tier

Facility Type	Facilities (n)	% of Total	Week 12 Compliance
Primary Health Centre (PHC)	47	49.0%	86.2%
Secondary / General Hospital	28	29.2%	93.1%
Tertiary / Teaching Hospital	12	12.5%	95.8%
Faith-Based Facility	9	9.4%	91.7%
Total	96	100%	89.6% (avg.)

Primary Health Centres (PHCs) comprised the largest facility category at 49.0% of the pilot sites (n = 47), followed by secondary/general hospitals (n = 28, 29.2%), tertiary/teaching hospitals (n = 12, 12.5%), and faith-based facilities (n = 9, 9.4%). The tertiary hospitals achieved the highest compliance rate of 95.8%, likely reflecting higher levels of digital literacy and stronger facility infrastructure. Faith-based facilities achieved compliance of 91.7% by week 12, while PHCs achieved 86.2%, a notable result given the documented

infrastructure and digital literacy challenges affecting many primary care facilities.

Facility Profile: Rural vs. Urban Location

The rural-urban status of the facilities was confirmed by program officers at the time of facility enrollment to the project based on the local government area (LGA) classification by the National Population Commission (NPC) and population density. See Table 4 for details.

Table 4. Facility Distribution and Compliance Rates by Location Type

Location	Facilities (n)	% of Total	Week 1 Compliance	Week 12 Compliance
Urban	41	42.7%	68.3%	93.4%
Peri-Urban	29	30.2%	63.0%	88.5%
Rural	26	27.1%	51.4%	83.7%
Total	96	100%	61.5%	89.6%

The rural-urban comparison highlights the effect of location on both initial and final compliance levels. Urban health facilities had a week-one compliance rate of 68.3%, higher than the 51.4% recorded among rural facilities. These findings are consistent with previous reports of inequalities in digital infrastructure, mobile network coverage, and health workers' digital literacy in urban and rural Nigeria. By week 12, however, rural facility compliance reached 83.7%, an increase of 32.3 percentage points from the initial 51.4% and the largest absolute increase recorded for any subgroup in the study. Compliance at peri-urban facilities reached 88.5% by week 12 and followed the system-wide compliance trend closely. Overall, these findings indicate that a simple, user-friendly Telegram interface can narrow

documentation-compliance gaps even where infrastructure constraints remain.

System Performance Under Connectivity Constraints

Another important aspect of the system evaluation was its performance in environments with unreliable or limited internet access. 22.9% or 22 of the 96 facility sites had been identified by program officers as high-connectivity-challenge (HCC) meaning facilities with poor mobile data coverage, with frequent network disruptions, or those that could only access the internet over 2G. Table 5 presents the performance of the mHealth system in standard-connectivity vs. high-connectivity-challenge facilities.

Table 5. System Performance Comparison: Standard vs. High-Connectivity-Challenge Facilities

Metric	Standard Connectivity (n = 74)	High-Connectivity-Challenge (n = 22)
Week 1 compliance rate	65.8%	47.3%
Week 12 compliance rate	92.4%	79.5%
Absolute compliance improvement (wk1–wk12)	+26.6 pp	+32.2 pp
Avg. upload confirmation response time	6.9 seconds	14.2 seconds
Upload retry events (timeout or failed send)	31 (2.1% of uploads)	89 (18.7% of uploads)
Successful upload completion after retry	29 / 31 (93.5%)	76 / 89 (85.4%)

A small number of retry events were seen due to timeouts or failures in sending the data to the object store. These events occur when a segment of a file cannot be successfully sent to the storage location and must be reattempted.

In this sub-section, we show how the connectivity analysis highlights both the strengths and the weaknesses of the proposed system in difficult network conditions. High-connectivity-challenge (HCC) facilities were those with initially lower dataflow compliance. In Week 1, dataflow compliance was only 47.3% for the 15 HCC facilities compared with 65.8% for the standard connectivity facilities. The average response time for the successful upload of a confirmation at the HCC facilities was 14.2 seconds per transaction, more than twice the 6.9 seconds for the standard connectivity facilities, owing to retransmissions of data packets and renegotiation for 2G connections. Despite the challenges, the HCC facilities showed the largest absolute increase in compliance across the pilot period, having

increased their compliance by 32.2 percentage points over the three months from Week 1 to Week 12, in contrast to an increase of 26.6 percentage points over the same period for standard facilities. In other words, it was the Health Centre (HCC) facilities that had endured the hardships of paper-based monitoring, and who therefore showed the greatest increase in compliance when they had surmounted the difficulties associated with connecting to the mobile technology and could therefore more effectively utilise the functionality provided by the system. The retry success rate of 85.4% also gives a good indication of the level of health worker commitment to compliance in the HCC facilities. These performance comparisons are summarised in

Figure 3 while the distribution of upload response times by connectivity category is shown in

Upload Confirmation Response

Figure 4, and compliance rates by facility tier across Week 1 and Week 12 are presented in

Hospital

Figure 5.

We find that week-12 compliance rates in HCCs were 79.5% (95% CI: 78.2%–80.8%)—lower than the 92.4% (95% CI: 92.1%–92.7%) achieved in standard-connectivity facilities—nevertheless representing a substantial

improvement relative to the pre-deployment manual monitoring baseline and serving to validate the efficacy of the low-bandwidth, data-efficient approach realized through file payload compression and asynchronous message transmission achieved in Telegram. These results further solidify the rationale in favour of integrating offline-capable upload buffering into future versions of the system, which would enable health center staff to initiate an upload transaction even when no connection is available, leaving the completion of the transaction up to the system when a connection is subsequently established.

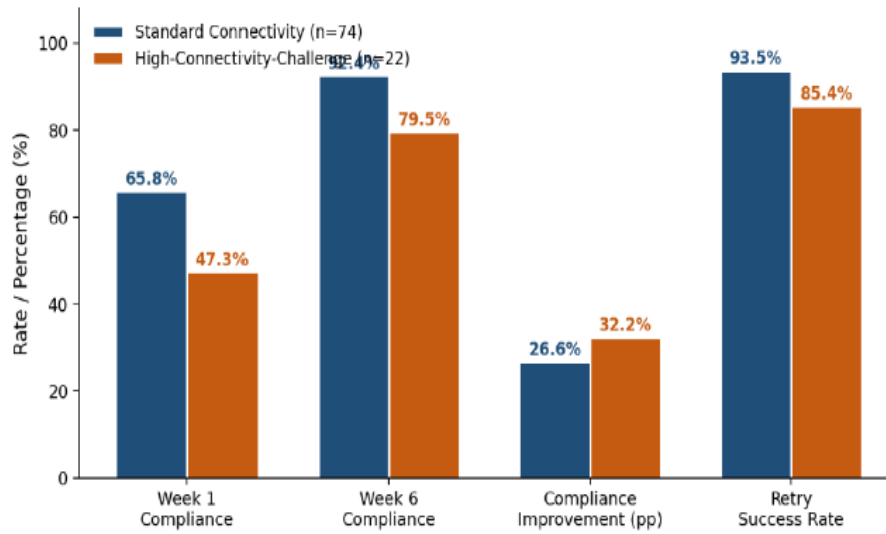


Figure 3. System Performance — Standard vs. High-Connectivity-Challenge Facilities

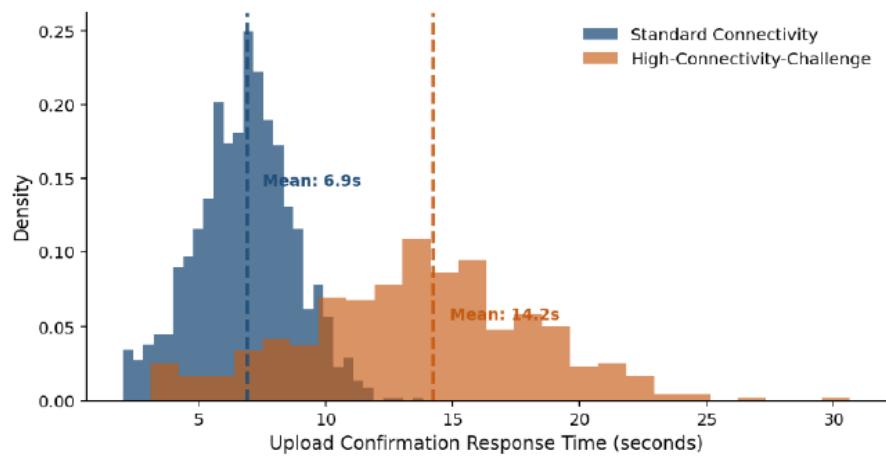


Figure 4. Distribution of Upload Response Times by Connectivity Category

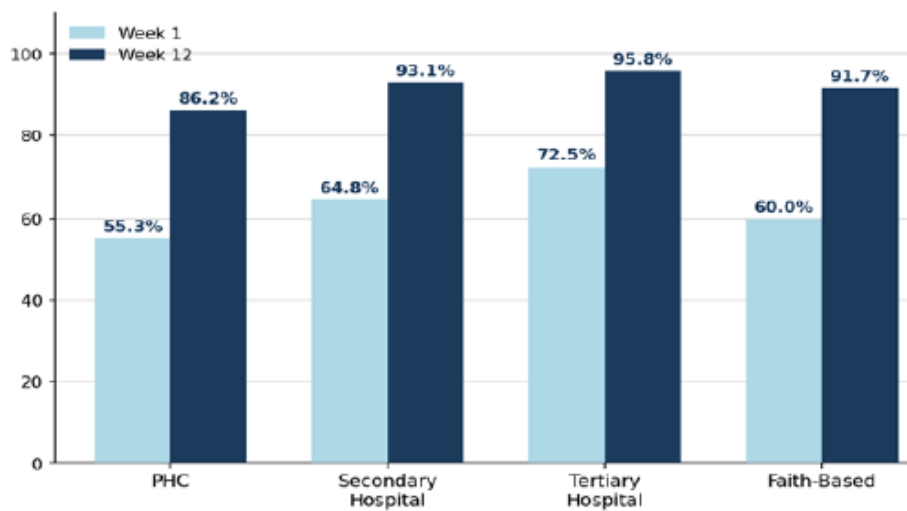


Figure 5. Compliance Rate by Facility Tier — Week 1 vs. Week 12

Documentation Compliance Trends

As highlighted earlier, the frequency of daily documentation was recorded over the 12-week pilot, and compliance increased throughout the duration of the trial. In the first week, compliance was 61.5%, compared with 89.6% by week 12. Therefore, compliance rose by 28.1 percentage points over the 12-week period. This pattern is consistent with mHealth implementation literature, which notes that automatic data collection systems often show compliance increases after an initial transition phase as staff adapt to a new workflow [19, 20].

Based on the learning curve of the Tanga region, we deduce that the barrier for documentation of incidents in the pre-automation phase was not the health workers' willingness, but rather the fact that the reporting was difficult. As mentioned earlier, [5] found that health information systems in several African health systems were inaccessible, which was a determinant to health workers' adoption of digital surveillance tools.

Comparative Advantage of the Approach

Our proposed system have many advantages when compared to current documentation tools. For example, users have to have access to the web via a browser in DHIS2, which is something that is generally not available due to low bandwidth in most areas, even in urban areas. This will not be the case for our proposed Telegram interface. In addition, HealthComX has shown that the use of instant messaging apps for health applications is also a viable option in limited resource settings, and given the near 100 % adoption of our interface in all 96 facilities that participated in compliance in our documentation interface also proves that our system has a very good potential for increasing the compliance of the necessary health sector documentation.

The automation project primarily relied on the workflow automation tool n8n. Adding new validation rules, notifications, and database

queries could be implemented almost instantly without having to write code. No-code and mHealth implementation literature characterizes this type of platform as a rapid development approach that can support innovation in low-resource health systems [19, 20]. The PostgreSQL audit trail has made it possible to ensure that adherence to measurement, reporting, and documentation compliance can be verified and reported at any given time. National HMIS guidance identifies routine, auditable facility-level data as a core requirement for health information management, and the PostgreSQL audit trail satisfies this requirement [23].

Implications for AI-Powered Surveillance

The data on documentation compliance that is generated by the system also creates a source of primary data that can be used for other purposes in the future, in particular as part of AI-based surveillance activities. Recent AI surveillance literature emphasizes that high-quality, complete, and up-to-date data are necessary for these systems to realise their full surveillance capabilities [16, 17, 23]. The documentation collection and logging system described in this paper facilitates the automation of facility-record collection on a daily basis, thereby providing data infrastructure that can potentially support future AI-based disease surveillance activities.

Limitations

The following limitations were noted: The first limitation is related to the time period of the pilot, which was only 12 weeks, which is unlikely to account for all seasonal periods of documentation or long term behavioral changes. The system currently allows any ZIP to be submitted regardless of the contents of the files, which means that a non-compliant looking document (one that would fail when opened) as well as an empty or non-document file can be submitted, as the content validation

has not yet been implemented. Another limitation of the system as it is currently deployed is that many facilities, particularly those serving rural populations, have inconsistent internet availability. Finally, because our evaluation was the first of its kind (to our knowledge), we did not have a comparable control facility population to which to compare our evaluation group, because the current system for monitoring manual submissions was based on a large binder and form for documenting entries. This precludes our being able to know what percent of facilities in a given population would increase rates of submittal under the current system, thus it is harder to understand whether an increase in facilities with submittals following the implementation of our new technology is an actual effect of the technology, versus an effect of having a more user-friendly monitoring tool.

Conclusion

This paper delves into the design, implementation and assessment of an AI-enabled automation system for documenting health facilities' adherence to DRHS documentation standards in Nigeria. It tackles one of the key challenges faced by using AI in health surveillance, namely that facility data need to be properly captured, submitted, and aligned to a common standard before they can be processed and analyzed. As recent AI surveillance and infection-control studies emphasize, the potential of AI surveillance activities heavily depends on high-quality, complete, and up-to-date data [17, 21, 22].

The present study implemented digital health record tools across 96 health facilities in Nigeria. The evaluation showed that the system was reliable, user-friendly, and associated with improved documentation compliance from 61.5% to 89.6% over the 12-week pilot period. These findings are consistent with mHealth implementation literature and reviews of digital health initiatives in low-resource settings, which show that digital technologies designed

around healthcare workers' routine practices and devices can address persistent health information management challenges [19, 20].

Content-level file validation, offline uploads, and integration with existing data stores such as DHIS2 should be validated in future work. Future development should also test AI analytics over structured data produced by the upload mechanism to support compliance-risk monitoring and predictive surveillance, consistent with recent literature on AI-enabled infection surveillance and health information systems [21-23].

Acknowledgement

The authors acknowledge the support of the health facility coordinators and program officers who participated in the pilot deployment of the system as well as the technical teams who provided the infrastructure required for the deployment and real time monitoring of the system during the validation exercise.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author's Contributions

Conceptualization, resources, validation, formal analysis, investigation, data curation, writing original draft, writing review & editing and funding were conceived by the corresponding author. The corresponding author also implemented the initial data hosting, the overall system workflow architecture, and data analysis. All authors have read and agreed to the published version of the manuscript.

Ethical Approval

This study did not involve any clinical data or human subjects. All data available from the system consists of metadata related to uploaded files (facility ID, date/time, file name) and does not include any personal health information

(PHI). All healthcare workers used the system as part of their normal workflow activities and did so on a voluntary basis.

Data Availability

The aggregated data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

- [1]. World Health Organization (WHO). Everybody's Business: Strengthening Health Systems to Improve Health Outcomes. WHO's Framework for Action. Geneva, Switzerland: World Health Organization; 2008.
- [2]. Olu O, Usman A, Woldetsadik S, Chamla D, Walker O. Lessons learnt from coordinating emergency health operations during the 2014 Ebola outbreak in West Africa. *Infectious Diseases of Poverty*. 2016;5(1):1-10.
- [3]. Mendes VIS, Mendes BMF, Moura RP, Lourenco IM, Oliveira MFA, Ng KL, et al. Harnessing artificial intelligence for enhanced public health surveillance: a narrative review. *Frontiers in Public Health*. 2025;13:1601151. Available from: <https://doi.org/10.3389/fpubh.2025.1601151>
- [4]. Tornimbene B, Leiva Rioja ZB, Brownstein J, et al. Harnessing the power of artificial intelligence for disease-surveillance purposes. *BMC Proceedings*. 2025;19(Suppl 4):7.
- [5]. Tshimula JM. Artificial intelligence for public health surveillance in Africa: Applications and opportunities. *PLOS Digital Health*. 2024;3(6):e0000512. Available from: <https://doi.org/10.1371/journal.pdig.0000512>
- [6]. Idahor CO. Infectious disease surveillance in the era of big data and AI. *Journal of Epidemiology and Public Health Reviews*. 2025;10(1):1-9.
- [7]. Villanueva-Miranda I, Xiao G, Xie Y. Artificial intelligence in early warning systems for infectious disease surveillance: a systematic review. *Frontiers in Public Health*. 2025;13:1609615. Available from: <https://doi.org/10.3389/fpubh.2025.1609615>

Facility-level operational audit logs are not publicly deposited because they contain program-specific metadata and facility identifiers.

Conflict of Interest

The authors declare that they have no conflict of interest.

- [8]. El Arab RA, Almoosa Z, Alkhunaizi M, Abuadas FH, Somerville J. Artificial intelligence in hospital infection prevention: an integrative review. *Frontiers in Public Health*. 2025;13:1547450. Available from: <https://doi.org/10.3389/fpubh.2025.1547450>
- [9]. Ferreira SAL. Hospital-acquired infection surveillance using machine learning. 2024.
- [10]. Radaelli D, Di Maria S, Jakovski Z, Alempijevic D, Al-Habash I, Concato M, Bolcato M, D'Errico S. Advancing patient safety: the future of artificial intelligence in mitigating healthcare-associated infections: a systematic review. *Healthcare*. 2024;12(19):1996. Available from: <https://doi.org/10.3390/healthcare12191996>
- [11]. Li JH. Artificial intelligence in infection surveillance: Data integration, applications and future directions. *Biomedical Journal*. 2026;49(2):100929. Available from: <https://doi.org/10.1016/j.bj.2025.100929>
- [12]. Alwakeel MM. AI-assisted real-time monitoring of infectious diseases in urban areas. *Mathematics*. 2025;13(12):1911. Available from: <https://doi.org/10.3390/math13121911>
- [13]. Vargas-Santiago M. A state-of-the-art review of artificial intelligence applications in healthcare: advances in diabetes, cancer, epidemiology, and mortality prediction. *Computers*. 2025;14(4):143. Available from: <https://doi.org/10.3390/computers14040143>
- [14]. Hevner AR, March ST, Park J, Ram S. Design science in information systems research. *MIS Quarterly*. 2004;28(1):75-105. Available from: <https://doi.org/10.2307/25148625>
- [15]. Dong C, Liu Y, Nie J, Zhang X, Yu F, Zhou Y. Artificial intelligence in infectious disease diagnostic technologies. *Diagnostics*.

2025;15(20):2602. Available from:
<https://doi.org/10.3390/diagnostics15202602>

[16]. Arora G, Joshi J, Mandal RS, Shrivastava N, Virmani R, Sethi T. Artificial intelligence in surveillance, diagnosis, drug discovery and vaccine development against COVID-19. *Pathogens*. 2021;10(8):1048. Available from:
<https://doi.org/10.3390/pathogens10081048>

[17]. Kim C, Austin R, Wurtz R, Delaney CW, Rajamani S. Applications of artificial intelligence in the control of infectious diseases in the post-COVID era: scoping review. *JMIR Nursing*. 2025;8:e84242. Available from: <https://doi.org/10.2196/84242>

[18]. Sundermann A, Chen J, Saul M, Babiker A, Kumar P, Griffith M, et al. Artificial intelligence enhances genomic surveillance in healthcare outbreak investigations. *Infection Control and Hospital Epidemiology*. 2026;47(2):170-4. Available from:
<https://doi.org/10.1017/ice.2025.10355>

[19]. Mehl G, Labrique A. Prioritizing integrated mHealth strategies for universal health coverage. *Science*. 2014;345(6202):1284-7. Available from:
<https://doi.org/10.1126/science.1258926>

[20]. Källander K, Tibenderana JK, Akpogheneta OJ, Strachan DL, Hill Z, ten Asbroek AH, Conteh L, Kirkwood BR, Meek SR. Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low- and middle-income countries. *Journal of Medical Internet Research*. 2013;15(1):e17. Available from:
<https://doi.org/10.2196/jmir.2130>

[21]. Gastaldi S. Advancing infection prevention and control through artificial intelligence. *Journal of Hospital Infection*. 2025;148:1-12. Available from:
<https://doi.org/10.1016/j.jhin.2025.01.004>

[22]. Barbati C, Viviani L, Vecchio R, Arzilli G, De Angelis L, Baglivo F, et al. Artificial intelligence use and performance in detecting and predicting healthcare-associated infections: a systematic review. *Artificial Intelligence in Medicine*. 2026;172:103321. Available from:
<https://doi.org/10.1016/j.artmed.2025.103321>

[23]. Federal Ministry of Health (FMOH). National Health Management Information System Policy and Guidelines. Abuja, Nigeria: Federal Ministry of Health; 2020.