JRIIE Journal of Research Innovation and Implications in Education
Center for Research Implications & Practice

Website: www.jriiejournal.com ISSN 2520-7504 (Online) Vol.10, Iss.2, 2026 (pp. 580 - 596)

Data Quality and Integration Challenges in Public Healthcare Information Systems: A Scoping Review

Matendo Didas[1], James Chambua[2] & Godfrey Justo[3]
[1,2,3]College of ICT, University of Dar es Salaam
[1]Centre for ICT, St. John's University of Tanzania, Dodoma, Tanzania
Email: matendodidas@gmail.com

***Abstract**: The quality and interoperability of data from public health information systems (PHIS) are critical for effective decision-making yet remain a global challenge. This scoping review maps literature published between 2020 and 2025 to identify key issues related to data integration (DI) and data quality (DQ), along with their impacts and proposed solutions. The review followed the Joanna Briggs Institute (JBI) methodology and analyzed studies from PubMed, Scopus, Web of Science, and IEEE Xplore. Out of 1,250 records, 48 studies met the inclusion criteria, focusing on real-world public health applications such as surveillance, routine monitoring, and outbreak response. The findings reveal four interconnected domains of challenges categorized as technical, data quality, organizational and governance, and ethical and legal challenges. These issues were further intensified during the COVID-19 pandemic, exposing weaknesses in public health data systems. The consequences of these challenges include reduced trust in public health institutions, inefficient resource allocation, and limited real-time surveillance capabilities. The review concludes that addressing DQ and DI challenges requires a holistic approach that integrates technical, organizational, and ethical dimensions. Key recommendations include sustained investment in data infrastructure, adoption of flexible and standardized data frameworks, and the establishment of collaborative governance models. Emphasis on ethical data use and equity is essential to ensure inclusive and effective PHIS. This study provides a foundation for guiding future research, policy, and practice toward more resilient health information systems.*

***Keywords:** Data Quality, Data Integration, Public Health Informatics, Health Information Systems, Interoperability, Data Governance, Scoping Review.*

How to cite this work (APA):

Matendo, D., Chambua, J. & Justo, G. (2026). Data Quality and Integration Challenges in Public Healthcare Information Systems: A Scoping Review. *Journal of Research Innovation and Implications in Education, 10*(2), 580 – 596. **https://doi.org/10.59765/pdcy7**

1. Introduction

The foundation of contemporary population health management is public health information systems (PHIS). In order to assist disease monitoring, outbreak investigation, health promotion, policy formation, and program assessment, they comprise a broad range of systems intended to gather, process, analyze, and distribute health-related data (World Health Organization, 2025). These systems' ultimate objective is to produce timely, accurate, and useful intelligence that helps public health professionals safeguard and enhance community health.

Data from vital statistics registries, laboratory information management systems (LIMS), electronic health records (EHRs), environmental monitoring systems, and syndromic surveillance platforms should ideally flow smoothly into a single, analytical environment. Public health organizations would be able to precisely assess the effects of treatments, identify at-risk populations, allocate resources effectively, and track disease trends in real-time thanks to this integrated data environment. This exact capacity to integrate and utilize a variety of data sources is the foundation of the public health "learning health system" concept (McGinnis et al., 2021).

But reality frequently fails to live up to this ideal. Data integration (DI) and data quality (DQ) are two major issues that public health organizations frequently face (Mashoufi et al., 2023). The term "data quality" (DQ) refers to the attributes of data that establish its dependability and suitability for use; these attributes are frequently evaluated across dimensions such as correctness, completeness, timeliness, consistency, and validity (Wiley, 2022).

Interoperability, or the capacity of various systems to interchange and utilize information, is crucial for DI, which entails the technical and procedural processes of merging data from several sources to produce a cohesive picture (Braunstein & Braunstein, 2022). Due mostly to the COVID-19 pandemic, the years 2020–2025 have been a crucial stress test for PHIS worldwide. Deeply ingrained flaws in the public health data infrastructure were revealed by the pressing requirement for quick data on cases, testing, contact tracing, hospitalizations, and immunizations. Many nations' pandemic response efforts were hindered by problems like inadequate demographic data, conflicting data formats between hospitals and health departments, and delayed case reporting (Tucker & Wang, 2021). Additionally, the quick development of new data sources, such as wastewater surveillance and wearable device data, has created new challenges as well as opportunities for integration into conventional PHIS.

A thorough synthesis of literature from this crucial age is scarce, despite the fact that many individual studies and publications have highlighted particular DQ and DI problems. To map the quickly changing body of data, pinpoint important ideas and difficulties, and draw attention to research gaps, a scoping review is a suitable methodology (Munn et al., 2018). This review aims to give a general picture of the situation rather than evaluate the quality of the evidence. Thus, the following objectives served as the basis for this scoping review:

1. To systematically map the literature published between 2020 and 2025 concerning DQ and DI challenges in public health information systems.
2. To identify and categorize the primary types of DQ and DI challenges reported.
3. To synthesize the documented consequences of these challenges on public health functions.
4. To provide recommendations for the challenges from literature.

For public health executives, informatics experts, and legislators entrusted with bolstering the data underpinnings of public health practice in the post-pandemic period, this review offers insightful information. The study is structured as follows: The next section provides a brief relevant literature to the study, sections three 3) and three (3) cover the methodology, results, and discussion of the study, respectively. The last section five (5) provides the conclusion and recommendations of the study.

2. Literature Review

2.1. Overview of Public Health Information Systems and Interoperability

Although the usage of public health information systems (PHIS) has greatly increased due to the quick digitization of healthcare, issues with data integration (DI) and quality (DQ) still exist. Large amounts of health data are produced, but their efficient utilization is limited by disjointed systems and poor interoperability, according to existing literature. Although interoperability is widely acknowledged as crucial for coordinated healthcare delivery, it is still not well implemented in many settings (Costa et al., 2025; Torab-Miandoab et al., 2023). Interoperability is defined as the ability of systems to exchange and meaningfully use data. This restriction limits PHIS's ability to facilitate prompt, evidence-based decision-making.

2.3 Overview of DQ and DI challenges

Technical Challenges in Data Integration: The persistence of technical barriers to DI is one recurrent theme in literature. These include insufficient semantic compatibility, inconsistent system architectures, and a lack of standardized data formats. Inconsistencies in data interpretation arise from heterogeneous systems' frequent reliance on several coding standards, including ICD, SNOMED, and LOINC (Ait Abdelouahid et al., 2023). Even though interoperability standards like HL7 FHIR provide potential answers, their adoption is patchy, particularly in environments with limited resources.

Clinical and public health data integration is, therefore, limited by the rarity of seamless data sharing between systems (El-Yafouri & Klieb, 2025).

Data Quality Dimensions and Challenges: Accuracy, completeness, consistency, and timeliness are important aspects of data quality that are well-documented in PHIS research. Errors in human data entry, duplication, missing information, and delayed reporting are frequently blamed for poor DQ. Reduced adherence to data management rules is a result of these difficulties, which are made worse by growing data volumes and human capacity constraints (Saini et al., 2022). As a result, clinical decision-making, disease surveillance systems, and public health measures are all compromised by low-quality data.

Organizational and Governance Barriers: Organizational and governance issues, in addition to technical ones, are crucial in determining DQ and DI results. Effective data management is sometimes hampered by fragmented data silos, inadequate financing, and unclear governance frameworks. These difficulties are especially noticeable in low- and middle-income countries (LMICs), where institutional capacity and digital health infrastructure are still scarce (Jayathissa & Hewapathirana, 2023). Attempts to synchronize data across systems are made more difficult by the lack of unified policies and standards.

Ethical and Legal Considerations: Legal and ethical problems are becoming more widely recognized as critical to DQ and DI discussions. Data sharing and integration activities may be impeded by concerns about data security, privacy, and consent management. The literature emphasizes the importance of strong regulatory frameworks that find a compromise between protecting individual rights and ensuring data accessibility. Fostering trust among stakeholders, particularly in large-scale public health projects, is critical for assuring compliance and encouraging effective use of health data.

2.4 Synthesis and Research Gap

Overall, the literature demonstrates that DQ and DI challenges in PHIS are multidimensional, encompassing technical, organizational, and socio-ethical domains. While technological innovations provide potential solutions, their effectiveness depends on strong governance, sustained investment, and context-sensitive implementation. Notably, most existing frameworks are developed in high-resource settings, with limited applicability to LMIC contexts. This highlights a critical gap and underscores the need for context-aware approaches to improving data quality and integration in public healthcare systems.

3. Methodology

This scoping review was conducted in compliance with the Joanna Briggs Institute (JBI) methodology for scoping reviews (Peters et al., 2020; Tricco et al., 2018) and the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) (Figure 1) checklist, which included components like eligibility criteria, search strategy, research selection, data graphing, and data synthesis.

Eligibility Criteria: The eligibility requirements primarily focused on studies that examined public health information systems at the local, state, federal, or international levels, such as those for vital statistics, immunization registries, chronic disease monitoring, infectious disease and syndromic surveillance, and outbreak management. In addition to data quality issues like timeliness, completeness, and accuracy, studies had to specifically address data integration and interoperability issues between systems, such as between electronic health records and public health agencies or across laboratory systems. Whitepapers from reputable health organizations like the CDC and WHO, conference papers, government reports, and peer-reviewed journal publications were all deemed acceptable sources. Additionally, taken into account were editorials and commentary that provide specific, fact-based analysis of these concerns.

Search Strategy: A comprehensive search strategy was developed and put into action in September 2025. The electronic databases that were searched were IEEE Xplore, PubMed, Scopus, and Web of Science. The search approach covered three topics: (1) "public health"; (2) "information systems" or "data systems"; and (3) "data quality" or "data integration" or "interoperability." Keywords and regulated vocabulary terms, like PubMed's Medical Subject Headings (MeSH), were combined. The search was limited to English-language articles published between 2020 and 2025. The search strings that combine keywords associated with the subject matter are summarized in Table 1.

Table 1: Search Strings and Boolean Operators

Search strings combined keywords
"Public Health" [Mesh] OR "public health" OR "population health") AND ("Information Systems" [Mesh] OR "health information systems" OR "public health surveillance" OR "electronic health records") AND ("Data Accuracy" [Mesh] OR "data quality" OR "data integration" OR "interoperability" OR "data governance") AND ("2020/01/01" [Published Date]: "2025/12/30" [Search strings combined keywords]

Study Selection: The research selection process was managed using Covidence systematic review software, which made it simpler to input, deduplicate, and screen all obtained records. There were two stages to the screening process. Two unbiased reviewers first assessed each record's titles and abstracts in compliance with the qualifying criteria. Second, the same reviewers gathered and carefully assessed complete texts of publications deemed potentially relevant using the predefined inclusion criteria. Any disagreements at either level were resolved by discussion, and when necessary, a third reviewer was consulted.

Data Charting: To direct the evaluation, a consistent data capture form was created and tested. Key details were recorded for every included study, starting with fundamental study attributes such as the authors, year of publication, title, country of origin, and study design. Whether the study concentrated on topics like COVID-19 surveillance, immunization initiatives, or chronic disease monitoring, the form also recorded the public health background. The exact data quality issues that were mentioned were also included in the extracted data, especially those pertaining to timeliness, accuracy, completeness, and other pertinent aspects. The form also recorded difficulties with data integration, such as organizational, semantic, and technical obstacles. The reported effects of these data issues were recorded, as were any advice or solutions put forth by the studies.

Data Synthesis: Qualitative content analysis was used to examine the retrieved data. Inductive coding was used to classify challenges and suggestions into emergent themes and categories. Tables and figures that show the frequency and connections between various concepts are used to support the conclusions, which are presented in a narrative synthesis.

4. Results and Discussion

The first search of the database produced 1,250 publishing records. Title and abstract screening were applied to 945 records after duplicates were eliminated. After the eligibility of 145 full-text publications was evaluated, 48 studies were found to satisfy all inclusion requirements. The research selection procedure is shown in the PRISMA flow diagram in Figure 1, and Table 2 provides a summary of some of the representative publications that were evaluated.

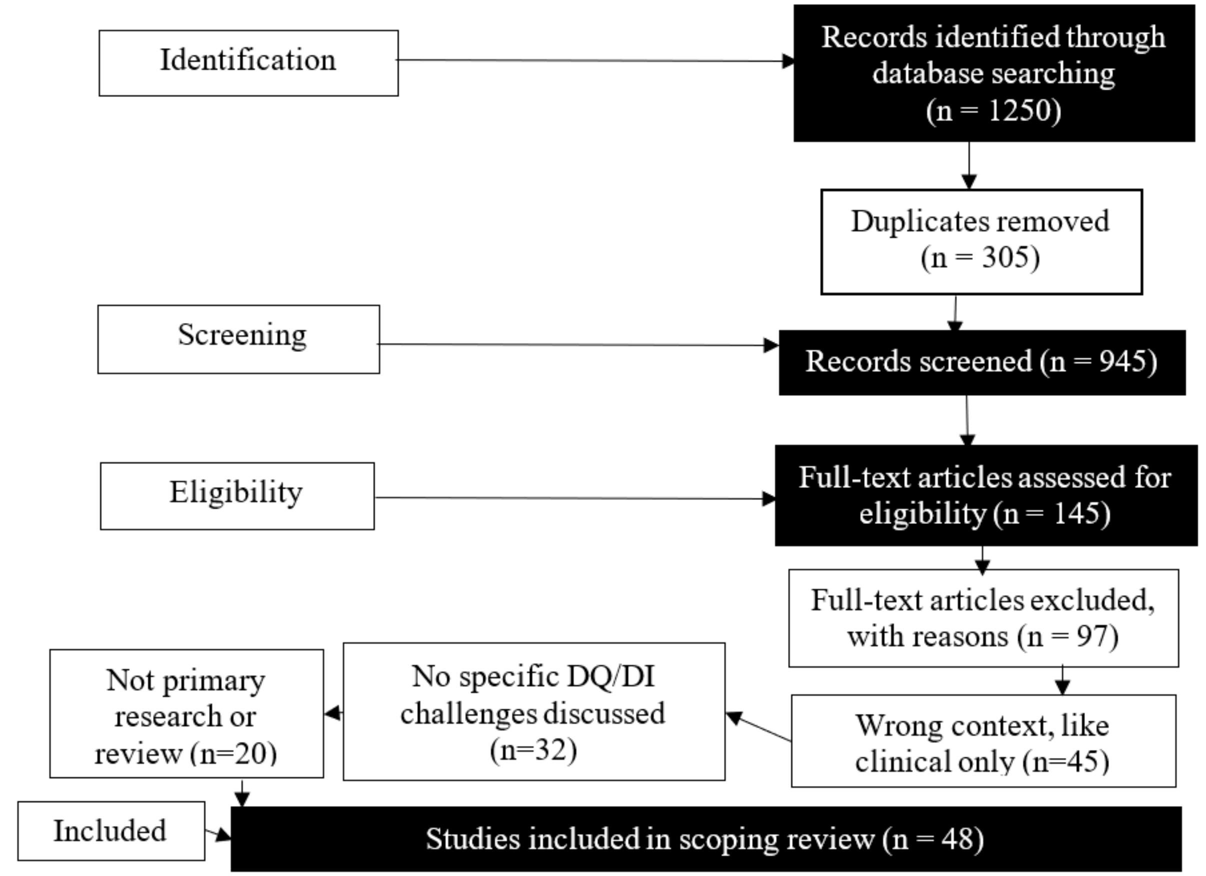

Figure 1: PRISMA Flow Representation of Study Selection

The majority (n=35) of the 48 included research focused on or were significantly impacted by the COVID-19 pandemic response, representing a wide range of geographic contexts and public health areas.

4.1 Categorization of Challenges

Technical, DQ, Organizational and Governance, and Ethical and Legal are the four key domains that emerged from the examination of the available literature. Figure 2 summarizes these problems.

Technical challenges: Technical challenges were the most frequently cited barrier to effective data integration. Since many studies (n=32) reported the continued usage of proprietary or heterogeneous structures across source systems, the lack of standardized data formats was a major issue. According to Chen et al. (2014) and Bakhrushin et al. (2020), hospitals that relied on different EHR suppliers often exported COVID-19 case data in incompatible formats that needed extensive manual transformation before being integrated into state surveillance platforms. Additionally, incompatible system architectures posed significant challenges because legacy public health systems based on outdated technologies were often unable to communicate with modern, cloud-based platforms used by labs or healthcare providers, leading to persistent data flow bottlenecks (Chitrakar, 2025; Adeshina, 2025). Even after syntactic interoperability was achieved, inconsistent definitions of important concepts, such as what constituted a COVID-19 death, and disparate coding systems, such as Logical Observation Identifiers Names and Codes (LOINC) versus local laboratory codes, hindered meaningful data integration and analysis (Ejaz, 2025).

Data quality challenges: The frequent reports of certain areas of data quality being insufficient weakened the value of integrated data. Real-time response efforts were hampered, especially in the early phases of the COVID-19 pandemic, by latency of days or even weeks between a positive test result and its input into national surveillance systems (Syed et al., 2023; Vasudevan et al., 2022). Studies show that punctuality has become a key issue. Completeness was another common issue. Critical demographic variables, including race, ethnicity, and occupation, as well as clinical fields like symptoms and vaccination status, were frequently missing, which hindered health equity analyses and made it more difficult to understand illness patterns (Spangler et al., 2023; Alofaysan et al., 2014). Duplicate records and unreliable epidemiological outputs were the result of inconsistent data entry, including typographical errors, inconsistent address formats, and incorrect laboratory results or case classifications; these problems were also commonly reported (Zhang et al., 2024a; Bernardi et al., 2023)

Organizational and governance challenges: these "soft" problems were frequently discovered to be the underlying causes of both technical and data-quality problems. Data silos, which resulted in vertical, disease-specific information systems like those for COVID-19, TB, and HIV that operate separately and are impossible to communicate, were first caused by fragmented funding and organizational structures (Arigbabu et al., 2024). Second, insufficient financing and personnel capabilities were a recurrent theme in numerous studies (n=28). These studies revealed ongoing underfunding of public health data infrastructure and a shortage of trained informatics personnel; as a result, many health departments lacked the funds required to implement complex integration initiatives or modernize systems (Ambalavanan et al., 2025; Alabi, 2023). Third, unclear policies regarding data ownership, stewardship, quality standards, and access controls were found to be a major barrier to the development of data-sharing agreements between organizations (Nazuli et al., 2025; Setiawan et al., 2025).

Ethical and Legal Challenges: The primary causes of ethical and legal issues include worries about consent, security, and data privacy. When integrating data from several sources, there is a contradiction between the need for granular data to enable meaningful analysis and the obligation to preserve individual privacy because this increases the risk of re-identification and data breaches (Zhang et al., 2024b). Integration efforts are further complicated by the legal uncertainty surrounding the sharing of patient information with public health agencies, especially for purposes other than immediate notifiable disease reporting (Nwankwo, 2025; Corrales Compagnucci & Fenwick, 2025).

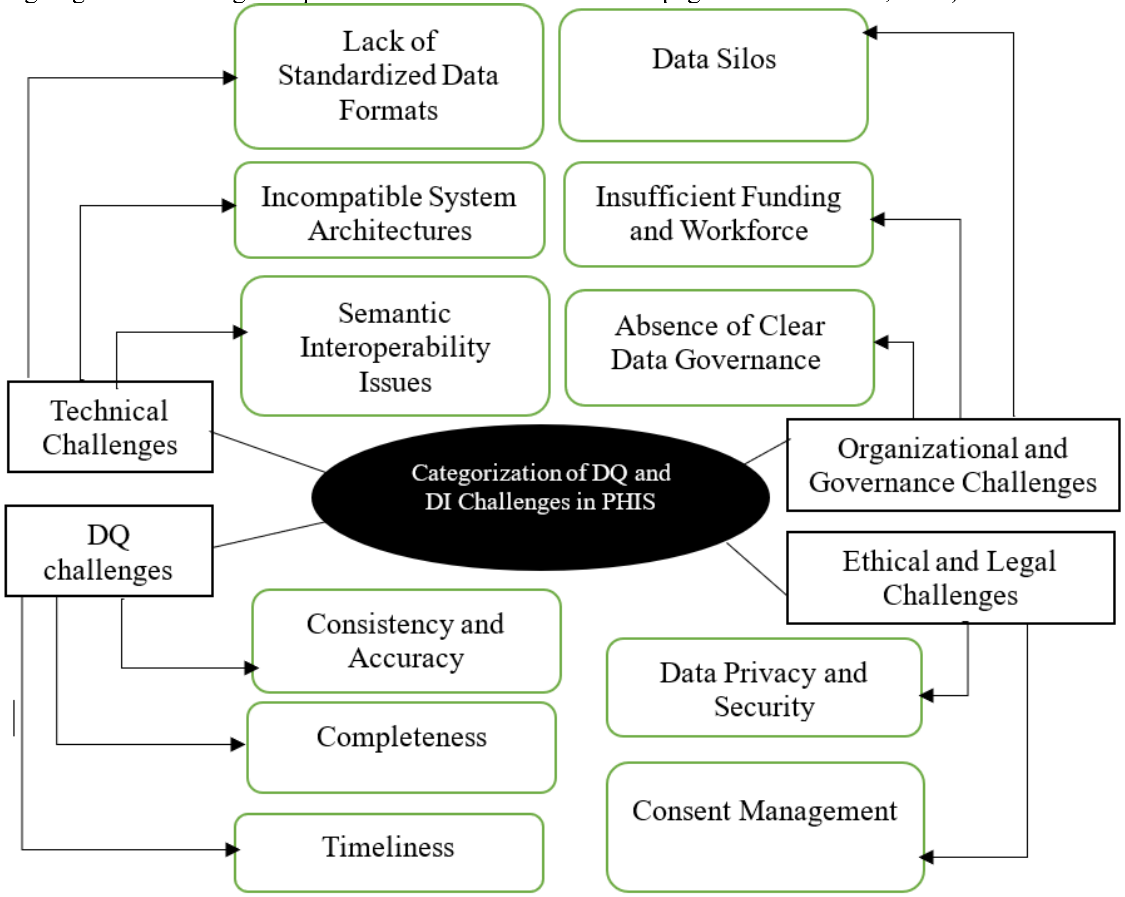

Figure 2: Categorization of DQ and DI Challenges in PHIS

3.2 Consequences of DQ and DI Challenges

The reviewed literature reveals several significant consequences arising from problems with data integration and quality, as seen in Figure 3. For example, situational awareness was hampered by fragmented datasets and low timeliness, forcing public health professionals to make critical decisions based on inadequate or outdated epidemic data (An et al., 2025; Sukumar et al., 2015). The deployment of testing kits, vaccines, and medical personnel was inefficient since it was challenging to identify hotspots and vulnerable individuals due to the lack of integrated data (Kim & Min, 2025; Bernardi et al., 2023). Additionally, diminished demographic completeness eventually hindered efforts to promote health equity among racial, ethnic, and socioeconomic groups by hiding disparities in disease burden and access to care (Wang et al., 2023). Confusion and a drop in public trust in health authorities were also caused by contradictory or inconsistent official data reporting during the pandemic (Bagyendera et al., 2024; Saini et al., 2022).

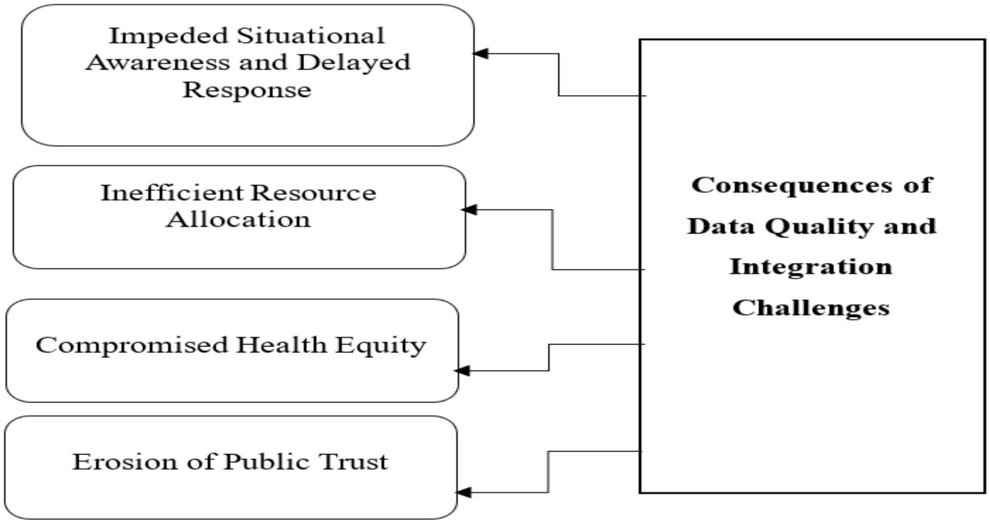

Figure 3: Consequences of Data Quality and Integration Challenges

Table 2: Representative Summary of Reviewed Literature (2020-2025)

Public Health Context	Key DQ Challenges	Key DI Challenges	Key Recommendations
Cross-organizational data sharing across public health agencies, laboratories, and healthcare providers	Typographical mistakes, inconsistent address formats, and incorrect classification are examples of inconsistency and inaccuracy that result in duplication and untrustworthy outputs (Zhang et al., 2024c; Bernardi et al., 2023)	Problems with semantic interoperability include disparate case definitions and inconsistent coding schemes, such as LOINC against local codes (Ejaz, 2025)	Create standardized case definitions, validation guidelines, and shared data dictionaries that are backed by governance organizations.
National and subnational public health governance and policy environments	Uncertain ownership, management, and accountability for data standards result in inconsistent data quality (Nazuli et al., 2025; Setiawan et al., 2023)	Data access, stewardship, and sharing agreements are unclear due to inadequate data governance frameworks (Nazuli et al., 2025; Setiawan et al., 2023)	Create official frameworks for data governance that include responsibilities, stewardship, quality requirements, and access restrictions.
Public health surveillance during infectious disease outbreaks, for example,	Inadequate promptness, which limits real-time reaction, with days or weeks passing between test	EHRs' disparate proprietary architectures and lack of defined data formats necessitate significant manual	To increase timeliness and decrease human processing, implement interoperable data

COVID-19, requires real-time reporting from hospitals, laboratories, and health departments.	findings and surveillance reporting (Syed et al., 2023; Vasudevan et al., 2022).	transformation (Chen et al., 2023; Bakhrushin et al., 2020)	standards such as HL7/FHIR and LOINC, and automate reporting pipelines
Multi-disease public health programs, for example, HIV, TB, and COVID-19, are operating within fragmented health systems.	Underinvestment and a lack of informatics capacity have resulted in persistent quality gaps that hinder the capacity to clean, validate, and manage data (Boppiniti, 2018; Alofaysan et al., 2014).	Vertical, disease-specific systems that lack communication can result in data silos (Arigbabu et al., 2024; Alabi, 2023)	Increase ongoing support for public health informatics workforce development and incorporate disease surveillance platforms.
Health Equity Analysis (Thompson et al., 2024)	Wholeness (race/ethnicity)	N/A	Require the gathering of standardized demographic information
Multi-source Data Integration	N/A	Security and privacy (Zhang et al., 2024b)	Make use of privacy-enhancing technologies (PETs) such as differential privacy.
Population-level epidemiological analysis and situational awareness	Equity studies are limited by incomplete data, particularly missing clinical variables (symptoms, vaccination status) and demographic variables (race, ethnicity, occupation) (Spangler et al., 2023; Alofaysan et al., 2014)	Modern cloud-based platforms are incompatible with legacy public health systems due to incompatible system architectures (Chitrakar, 2025; Adeshina, 2025)	Invest in scalable, cloud-compatible public health data infrastructure and update outdated systems.
Public communication and trust during health emergencies	Decreased credibility as a result of conflicting or inconsistent government data reporting (Moses et al., 2024; Saini et al., 2022)	Decreased credibility as a result of conflicting or inconsistent government data reporting (Moses et al., 2024; Saini et al., 2022)	To increase consistency and public trust, bolster coordinated reporting systems and openness
Use of individual-level health data for public health analytics and decision-making	Health equity initiatives are hampered by the inability to examine inequities because of incomplete or subpar demographic data (Moses et al., 2024)	Legal and ethical obstacles, such as re-identification issues, privacy threats, and ambiguous consent for the use of secondary data (Zhang et al., 2024a; Nwankwo, 2025)	Adopt strong de-identification, privacy-by-design strategies, and explicit legislative guidelines for the use of public health data.
Routine HMIS / facility & district HMIS data assessment	Limited timeliness; conflicting registrations versus reports; low completeness and inconsistent accuracy across indications	Parallel digital and paper systems; disjointed, non-interoperable data sources (Rumisha et al., 2020)	Boost HMIS training, do regular DQA, align registers, enhance feedback loops, and employ
Data quality assessment methods for public health	Measurement bias and inconsistent DQ definitions and measures across studies	Few validated integration measures and a lack of standard ways to evaluate integration (Shen et al., 2024)	Adopt standardized DQ frameworks, integrate techniques (audits, external validation), and consistently report DQ data.
Interventions to improve RHIS data quality and use	Inadequate supervision, human factors (skills, workload), and infrastructure that affect accuracy and completeness	System fragmentation has been observed to impede successful integration (Lemma et al., 2020)	Interventions with many components (training, supervision, feedback, and digital tools) enhanced DQ/use.
Quality assessment of maternal & child health	Variability in accuracy, internal inconsistency, and underreporting	Inadequate connections between district HMIS, facility registers, and higher-level systems	Consistent data validation, focused capacity growth, and

data in HMIS (Ouedraogo et al. 2019)			standardization of data gathering instruments
Routine HMIS data quality evaluation	Systematic quality difficulties with numerous indicators; denominator problems	Inconsistent aggregation procedures and gaps between paper and electronics (Adane et al., 2021)	Enhance source registers and offer instruction on DQA procedures and indicator definitions.
Composite DQ scoring & DHIS2 data evaluation	Variability in consistency and completeness; composite measurements are required for cross-indicator comparisons	DQ is impacted by integration discrepancies in the aggregated feeds that DHIS2 receives from numerous upstream systems (Odeny et al., 2023)	Utilize composite DQ scores, standardize upstream data streams, and automate DQA on DHIS2.
EHR implementation impact on data quality	Workflow fit and user training have an impact on data accuracy and completeness.	Double input is caused by incompatibilities between national reporting systems and EHRs (Fraser et al., 2024)	Implement exchange standards, offer ongoing training, and match EHR design to workflows (FHIR/HL7)
EMR interoperability in referral hospitals (Mkayula et al., 2022)	Incomplete capture and inconsistent coding have an impact on data quality	There were several EMRs with different interfaces, and no single patient identification system.	Adoption of standards and unique IDs, partner collaboration, and national interoperability policies
Community & district health information systems	Problems with data timeliness, completeness, and downstream utilization	Inconsistent data models, fragmented CBHIS components, and little cross-sector connections (Kuvuna et al., 2024)	Invest in interoperable architectures, standardize information, and bolster governance.
Health data digitalization policy/roadmaps (World Health Organization, 2025)	Inadequate digitalization results in inadequate data interpretation and completeness	Standards and governance are frequently lacking, and many implementations are compartmentalized.	Governance structures, interoperability standards, national digital health plans, and capacity building
HIS performance assessment	Inadequate data validation, staff ignorance, and limited resources	There is minimal patient-level integration in systems intended for aggregated reporting (Teklegiorgis et al., 2016)	Prioritize staff skill development, implement validation checks, and make sure the right tools are available.
Factors influencing data quality in routine HIS (Morris et al., 2024)	Data quality is harmed by organizational, behavioral, and technical variables (such as employee turnover).	Inability of NGOs and governmental entities to use interoperable systems	Invest in user support, align NGO reporting, and strengthen governance
Improving routine maternal and newborn data	Data errors brought on by inadequate oversight and a lack of local data review	Disjointed data sources (facility reports versus registries), complex integration (Bhattacharya et al., 2020)	Integrated strategy with a district focus (training, mentoring, plus review) enhanced DQ.
EMR impact on reporting completeness and timeliness (Ngugi et al., 2022)	EMRs increased timeliness, but depending on workflows, their accuracy and completeness varied	Consistent automated reporting to DHIS2 was hampered by EMR heterogeneity	Invest in ETL, standardize EMR-to-DHIS2 pipelines, and keep an eye on data flow quality.
MCH data quality assessment via DHIS2	Completeness and timeliness are generally good; there are some facility-level errors.	Difficulties in integrating DHIS2, e-tracker, and facility registers consistently (Lasim et al., 2022). Consistent automated reporting to DHIS2 was hampered by EMR heterogeneity	Boost source documentation, oversight, and regular e-tracker and DHIS2 reconciliation.

Perceived accuracy and use of DHIS2 data for RMNCAH+N	Stakeholders believed that use was limited by gaps in accuracy and synthesis skills.	Stakeholders highlighted inconsistencies and a lack of provenance in DHIS2 aggregates from several upstream systems (Oware et al., 2025)	Invest in provenance documentation, data synthesis training, and stakeholder coordination.
Barriers to healthcare data quality in urban public facilities	Human factors, inadequate data management abilities, and a weak DQA culture	Multiple unconnected digital technologies; integration is not prioritized (Tolera et al., 2024)	Encourage a culture of system rationalization, focused training, and high-quality data.
Routine HIS data quality formative survey (Gobena et al., 2022)	Differences in content completeness and accuracy between facilities	Inconsistent system flows and a lack of defined aggregation procedures	Bolster DQA procedures, HIS training, and fundamental computation capabilities
Routine use of DHIS2 data in LMICs (Byrne & Sæbø, 2022)	Staff capacity and perceived dependability limit data utilization.	DHIS2 relies on the quality of upstream data collection; integration issues arise when systems export various formats	Enhance the culture of data use, standardize indicators, and make upstream system quality investments.
Interoperable DHIS2-LMIS (logistics) use cases (Rogers, 2023)	Inconsistencies in data between aggregated reports and facility stock records	Demonstrated the practical challenges of combining LMIS with DHIS2 (vocabularies, mapping)	Clearly define the mapping, use middleware for deterministic linkage, and share information and IDs
Strengthening immunization RI data quality (Noiki et al., 2025)	Long-standing problems with duplicate reporting and the quality of immunization data	Inadequate program harmonization and many data reporting streams	improving the system, integrating RI reporting into a single DHIS2 instance, and providing mentorship
Data quality and use in primary health care (immunization)	Double-input, inaccurate registrations, and gaps in data entry skills	Inconsistency between sources was caused by paper-digital double entry (Eze et al., 2023)	Minimize duplicate entries, digitize at the time of capture, and increase oversight
District integrated approach to improve routine data (Avan et al., 2024)	When interventions included technical and governance efforts, data quality improved.	When district systems were coordinated and unique identities were utilized, integration improved	It is advised to use integrated district support packages that include governance, tools, and mentoring.
Community / District-level HIS data quality management	Widespread poor data quality, including inadequate data administration at the district and neighborhood levels, incompleteness, inaccuracy, unreliability, and tardiness	Integration between systems or between levels (facility, district) was either nonexistent or very limited because many systems are decentralized and paper-based (Ndabarora et al., 2014)	Establish robust HIMS management and feedback loops, invest in data management capabilities, and implement systematic HIS structures. Aim for data utilization rather than merely collecting.
Health information exchange (HIE) for digital health systems	Integration and policy were given more attention.	Limited interoperability between systems and jurisdictions, fragmented digital health systems, and a lack of standardized data sharing protocols (Mamuye et al., 2022)	At the national level, create and execute thorough, compatible technical standards; provide frameworks for data use, security, privacy, and governance; increase technological and human resource capacity.
Health systems strengthening via data-driven quality	Poor data reliability eroded confidence in data, and data quality	Lack of a single HIS, several data sources and reporting streams, and	Instead of using single fixes, use multicomponent interventions (data audits, feedback loops, capacity

improvement (QI) (Wagenaar et al., 2027)	and accessibility were frequently insufficient to support QI.	challenges combining data for cross-level decision-making	training, system strengthening); foster a culture of data use among employees at all levels.
Transition from paper-based to digital HIS; data validation & use for health action	Data quality does not immediately increase with digitalization; there are still problems with validation, trust, and completeness.	Poppe et al. (2021) and Collinson et al. (2021) reported that the integration of new digital platforms with legacy/paper systems frequently failed, resulting in fragmented workflows and poor module compatibility.	Before a broad rollout, prioritize data validation procedures, encourage feedback and the use of data for action, involve stakeholders in the design process, and make sure digital solutions match local workflows.
Integration of eHealth technologies with routine hospital work activities (Achieng & Ruhode, 2021)	Mismatches between technology design and actual workflows have an impact on data quality and can result in system underuse or misuse.	Lack of standardization, poor alignment with hospital procedures and human practices, and fragmentation and lack of cooperation across many eHealth apps	Before system design, do a practical evaluation of workflow fit; include end users; improve stakeholder coordination; and see HIS integration as a socio-technical shift rather than merely a technical implementation.
HIS interoperability in LMICs, review paper (Jayathissa & Hewapathirana, 2023)	Put more emphasis on integration.	Integration is hampered by inconsistent data formats and protocols, disjointed systems across vertical projects, a lack of resources, and data privacy and security issues.	Poor cross-system coordination, a lack of a uniform architecture, and numerous disparate health systems (EHRs, registries, vertical-program databases) that don't connect
Routine health information system (RHIS) data quality assessment (Adejumo & Criss, 2017)	Inconsistency between facility records and aggregated reports; inadequate data correctness and completeness at the facility level	Weak integration with higher-level reporting systems; inadequate infrastructure or a lack of standard operating procedures frequently causes data flow disruptions	To decrease disparities, improve facility-level data recording procedures, conduct frequent data audits, provide training and supervision, and establish clear reporting guidelines
Syntactic interoperability needs common data models for HIS	Not the main focus of data quality	Systems cannot successfully communicate structured data due to a lack of a shared data model or established data definitions, as well as a lack of syntactic interoperability (Moses et al., 2024)	Lack of a common standard, numerous heterogeneous HIS systems with disparate data schemas, and challenges connecting or sharing data between systems
Use of routine health information system (RHIS) data for research and evaluation (Hung et al., 2020)	Data flaws include missing data, inconsistent or untrustworthy data across facilities, and issues with representativeness and completeness for use in research.	Integration problems arise when integrating RHIS data from many systems or sources for study or assessment; harmonization is challenging due to a lack of metadata regarding data origin.	Encourage clear documentation of data gathering techniques and constraints; create strategies for handling incomplete data, such as cleaning and validation; and improve data consistency and completeness at the source to enable trustworthy secondary use.

This scoping review covered the literature from 2020 to 2025, a period that put public health data systems to the test globally. Despite technological advancements, the results continually demonstrate that PHIS is still hampered by a complex interplay of technical, data quality, organizational, and ethical challenges. The COVID-19 pandemic was not the cause of these problems; rather, it was a trigger that exposed and made worse systemic weaknesses that already existed. One important idea that came out of this review is the interconnectivity of challenges. For example, poor data quality in the form of inconsistencies (a DQ issue) is sustained by a lack of enforced data governance (an organizational issue), which is directly caused by the technical challenge of non-standardized data formats (a technical issue). Because of this interdependence, isolated solutions that focus on a single domain are probably going to fail. A comprehensive and well-coordinated approach is necessary for success.

A persistent issue that requires particular attention is semantic interoperability. While standards like HL7 FHIR eventually resolve syntactic interoperability (the ability to transfer data), the "last mile" problem of ensuring consistent semantics remains (Pandya, 2026). This highlights the need for a dual focus: adopting technology standards and creating collaborative communities of practice to agree upon and implement shared data models and value sets for public health. The analysis also highlights that, rather than being technological, the largest challenges are often socio-organizational. Data silos are frequently caused by a lack of trust across agencies, institutional goals, and disparate funding sources. To overcome these challenges, strong leadership, enduring political will, and investment in interjurisdictional relations are required. In this case, switching to robust data governance systems is essential. Such frameworks must be collaborative and incorporate all pertinent parties, such as public health organizations, labs, IT suppliers, and healthcare providers, in order to establish clear criteria for data sharing, quality, and usage.

Privacy and equity are becoming more and more important ethical concerns. When PHIS tries to incorporate more diverse and detailed data sources, the potential of algorithmic bias and privacy infringement rises (Zhang et al., 2024c). For example, the profession must aggressively adopt privacy-enhancing technology and integrate equity issues into the basic architecture of data systems by mandating the collection of extensive, standardized demographic data (Thompson et al., 2025).

Lastly, the assessment identifies a crucial disconnect between suggested fixes and their execution. Few studies addressed the sustainable finance and personnel development needed to execute and maintain these solutions, despite the fact that many suggested technological remedies like APIs or data lakes. The elephant in the room is the persistent underfunding of public health infrastructure; if it is not addressed, the cycle of crisis-exposure and short-term fixes will persist.

5. Conclusion and Recommendations

The literature on data quality and integration issues in public health information systems (PHIS) from 2020 to 2025 is thoroughly synthesized in this scoping study. The results verify that these issues are intricate, systemic, and seriously impair public health outcomes and performance. These flaws were further highlighted by the COVID-19 pandemic, which showed that robust and adaptable public health systems rely on excellent, interconnected data infrastructures. The review emphasizes the necessity of moving toward unified and interoperable public health data ecosystems from disjointed, disease-specific systems. More than just technological fixes are needed to bring about this change; ongoing funding, flexible data standards, cooperative governance frameworks, and a strong dedication to moral values and health justice are all necessary. Developing PHIS that are appropriate for the present period requires addressing the interplay of organizational, socio-ethical, and technical difficulties.

There are several limitations to this study. First, pertinent research from non-English-speaking cultures would have been left out if only English-language publications were included. Second, the database search ended in September 2025, which may have left out current advancements and gray literature in this quickly developing topic. Third, because it was a scoping review, the study concentrated on mapping the range of evidence rather than evaluating the methodological quality of the included papers. As a result, the frequency of difficulties that are reported could not accurately represent how serious they really are. Lastly, the underrepresentation of successful case studies on efficient data integration and quality improvement may have resulted from the focus on identifying obstacles. Important recommendations for efficient DQ and DI of PHIS are highlighted in this scoping review (Table 3).

Table 3: Recommendations of the study

S/N	Recommendation	Description
1	**Policy and Investment**	To move past temporary fixes and toward long-term system modernization, policymakers and funders should place a high priority on consistent and adaptable investments in public health data infrastructure. Interoperable, cloud-enabled technologies that facilitate scalability and resilience should be prioritized. To improve interoperability, regulatory regimes should require the use of standardized data formats like HL7 FHIR and terminologies like LOINC and SNOMED-CT.
2	**Workforce Development**	Dedicated funding is needed to strengthen the public health informatics workforce through targeted training, recruitment, and retention strategies. Building human capacity is critical to ensuring the effective implementation and maintenance of integrated data systems.
3	**Governance and Coordination**	Establishing cooperative, multi-stakeholder governance structures with distinct control over data standards, sharing procedures, and access guidelines is a good idea for public health organizations. Improving coordination between systems and stakeholders and reducing data silos are two benefits of strengthening governance.
4	**Data Quality Improvement**	Basic data quality procedures, such as automated validation at data entry points, regular data quality audits, and consistent data gathering procedures, should be given top priority by organizations. To promote equity-focused public health interventions, thorough and consistent demographic data should receive special emphasis.
5	**Technological Innovation**	Adopting API-first architectures can facilitate seamless and automated data exchange, reducing reliance on manual processes and improving system efficiency. Investments in modern integration technologies are essential for enabling real-time data sharing.
6	**Future Research and Practice**	Future research should focus on socio-technical and implementation science approaches to address organizational and governance barriers to data integration. There is also a need to develop practical ethical frameworks and tools that support privacy-preserving data use and algorithmic fairness. Additionally, documenting and disseminating best practices through detailed case studies will help build a shared evidence base to guide future advancements in PHIS.

References

Achieng, M., & Ruhode, E. (2021). EHealth Technologies Integration with Healthcare Work Activities in Public Hospitals: A Critical Realist Perspective. *arXiv preprint arXiv:2108.09734*.

Adane, A., Adege, T. M., Ahmed, M. M., Anteneh, H. A., Ayalew, E. S., Berhanu, D., ... & Janson, A. (2021). Routine health management information system data in Ethiopia: consistency, trends, and challenges. *Global health action*, *14*(1), 1868961.

Adejumo, A., & Criss, J. (2017). An assessment of data quality in routine health information systems in Oyo State, Nigeria. *University of the Western Cape*.

Adeshina, Y. T. (2025). Interoperable IT architectures enabling business analytics for predictive modeling in decentralized healthcare ecosystems. *Unpublished/White Paper or Technical Report*.

Ait Abdelouahid, R., Debauche, O., Mahmoudi, S., & Marzak, A. (2023). Literature review: clinical data interoperability models. *Information*, *14*(7), 364.

Alabi, M. (2023). Data governance and quality: Ensuring data reliability and trustworthiness. *ResearchGate*.

Alofaysan, S., Alhaqbani, B., Alseghayyir, R., & Omar, M. (2014, March). The significance of data governance in healthcare. In *Proceedings of the International Conference on Health Informatics*.

Ambalavanan, R., Snead, R. S., Marczika, J., Towett, G., Malioukis, A., & Mbogori-Kairichi, M. (2025). Challenges and strategies in building a foundational digital health data integration ecosystem: a systematic review and thematic synthesis. *Frontiers in health services*, *5*, 160068

An, D., Lim, M., & Lee, S. (2025). Challenges for data quality in the clinical data life cycle: systematic review. *Journal of Medical Internet Research*, *27*, e60709.

Arigbabu, A. T., Olaniyi, O. O., Adigwe, C. S., Adebiyi, O. O., & Ajayi, S. A. (2024). Data governance in AI-enabled healthcare systems: A case of the Project Nightingale. *Asian Journal of Research in Computer Science*, *17*(5), 85-107.

Avan, B. I., Dubale, M., Taye, G., Marchant, T., Persson, L. Å., & Schellenberg, J. (2024). Data-driven decision-making for district health management: a cluster-randomised study in 24 districts of Ethiopia. *BMJ Global Health*, *9*(2).

Bagyendera, M., Nabende, P., Godman, B., & Nabukenya, J. (2024). Contextualizing Syntactic Interoperability Data Standards for Health Information Exchange in Uganda's Public Healthcare System. In *ICT4AWE* (pp. 267-275).

Bakhrushin, V., Bakurova, A., Pasichnyk, M., Tereschenko, E., Pickl, S. W., Lytvynenko, V., ... & Sherstjuk, V. (2020). Risks of Data Inconsistency in Information Systems Used for Predicting the Development of Pandemics. In *CITRisk* (pp. 1-15).

Bernardi, F. A., Alves, D., Crepaldi, N., Yamada, D. B., Lima, V. C., & Rijo, R. (2023). Data quality in health research: an integrative literature review. *Journal of Medical Internet Research*, *25*, e41446.

Bhattacharya, A. A., Allen, E., Umar, N., Audu, A., Felix, H., Schellenberg, J., & Marchant, T. (2020). Improving the quality of routine maternal and newborn data captured in primary health facilities in Gombe State, Northeastern Nigeria: a before-and-after study. *BMJ open*, *10*(12), e038174.

Boppiniti, S. T. (2018). Unraveling the Complexities of Healthcare Data Governance-Strategies, Challenges, and Future Directions.

Braunstein, M. L., & Braunstein, M. L. (2022). *Health Informatics on FHIR: How HL7's API is transforming healthcare*. New York, NY: Springer.

Byrne, E., & Sæbø, J. I. (2022). Routine use of DHIS2 data: a scoping review. *BMC Health Services Research*, *22*(1), 1234.

Chen, H., Hailey, D., Wang, N., & Yu, P. (2014). A review of data quality assessment methods for public health information systems. *International journal of environmental research and public health*, *11*(5), 5170-5207.

Chitrakar, P. (2025). Cross-Platform Engineering and Enterprise Integration in Healthcare Information Systems: A Comprehensive Review.

Collinson, N., Saugene, Z., Braa, J., Sahay, S., & Mosse, E. (2021). Reflections, learnings, and proposed interventions on data validation and data use for action in health: A case of Mozambique. *arXiv preprint arXiv:2108.09731*.

Corrales Compagnucci, M., & Fenwick, M. (2025). A multidisciplinary perspective on cross-border health data transfers: privacy, risks and solutions. In *International transfers of health data: a global perspective* (pp. 1-15). Singapore: Springer Nature Singapore.

Costa, T., Borges-Tiago, T., Martins, F., & Tiago, F. (2025). System interoperability and data linkage in the era of health information management: A bibliometric analysis. *Health Information Management Journal*, *54*(3), 214-226.

Ejaz, U. (2025). Cross-Institutional Laboratory Data Sharing: Addressing Standardization in Public Health Emergencies.

El-Yafouri, R., & Klieb, L. (2025). A scoping review of electronic health records interoperability levels, expectations, approaches, and problems. *Health Informatics Journal*, *31*(4), 14604582251385986.

Eze, I. I., Ogbu, M., Ossai, E. N., Ekenna, A., Okoronkwo, I., & Onwujekwe, O. (2023). Data quality and use in primary health care: a case study of the immunization program in Ebonyi State, Nigeria. *Nigerian Journal of Clinical Practice*, *26*(Suppl 1), S143-S152.

Fraser, H. S., Mugisha, M., Bacher, I., Ngenzi, J. L., Seebregts, C., Umubyeyi, A., & Condo, J. (2024). Factors influencing data quality in electronic health record systems in 50 health facilities in Rwanda and the role of clinical alerts: cross-sectional observational study. *JMIR Public Health and Surveillance*, *10*(1), e49127.

Gobena, T., Shore, H., Berhanie, D., Kenay, A., Wondirad, Y., & Ayanle, M. (2022). Routine health information system data quality and associated factors in selected public health facilities of Jigjiga Woreda, Somali Regional State, Eastern Ethiopia. *The Ethiopian Journal of Health Development*, *36*(1).

Hung, Y. W., Hoxha, K., Irwin, B. R., Law, M. R., & Grépin, K. A. (2020). Using routine health information data for research in low-and middle-income countries: a systematic review. *BMC Health Services Research*, *20*(1), 790.

Jayathissa, P., & Hewapathirana, R. (2023). Enhancing interoperability among health information systems in low-and middle-income countries: a review of challenges and strategies. *arXiv preprint arXiv:2309.12326*.

Kim, S., & Min, W. K. (2025). Toward high-quality real-world laboratory data in the era of healthcare big data. *Annals of Laboratory Medicine*, *45*(1), 1-11.

Kuvuna, B., Nyanchoka, M., Guleid, F., Ogutu, M., Tsofa, B., & Nzinga, J. (2024). Community-based health information systems in Africa: a scoping review of data generation, utilization, and community empowerment. *Wellcome Open Research*, *9*, 485.

Lasim, O. U., Ansah, E. W., & Apaak, D. (2022). Maternal and child health data quality in health care facilities at the Cape Coast Metropolis, Ghana. *BMC Health Services Research*, *22*(1), 1102.

Lemma, S., Janson, A., Persson, L. Å., Wickremasinghe, D., & Källestål, C. (2020). Improving quality and use of routine health information system data in low-and middle-income countries: a scoping review. *PloS one*, *15*(10), e0239683.

Mamuye, A. L., Yilma, T. M., Abdulwahab, A., Broomhead, S., Zondo, P., Kyeng, M., ... & Tilahun, B. C. (2022). Health information exchange policy and standards for digital health systems in Africa: a systematic review. *PLOS Digital Health*, *1*(10), e0000118.

Mashoufi, M., Ayatollahi, H., Khorasani-Zavareh, D., & Boni, T. T. A. (2023). Data quality in health care: main concepts and assessment methodologies. *Methods of Information in Medicine*, *62*(01/02), 005-018.

McGinnis, J. M., Fineberg, H. V., & Dzau, V. J. (2021). Advancing the learning health system. *N Engl J Med*, *385*(1), 1-5.

Mkayula, N., Mbise, M., & Mahundi, M. (2022). Approaches towards interoperability of electronic medical records systems: a case of selected referral hospitals in Tanzania. *Journal of Health Informatics in Africa*, *9*(1), 1-13.

Morris, L. D., Nyongesa, M. W., & Sokiri, T. D. (2024). Factors influencing data quality in routine health information systems in Maridi county, South Sudan. *South African Journal of Information Management*, *26*(1), 1856.

Munn, Z., Peters, M. D., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, *18*(1), 143.

Nazuli, M. F., Walhidayah, I., Herawati, N. A., Yulianti, L. P., & Surendro, K. (2025). Data Governance Design for Optimization of Hospital Management Information System (SIM-RS) at ABC Regional Hospital. *Lontar Komputer: Jurnal Ilmiah Teknologi Informasi*, *16*(01), 27-40.

Ndabarora, E., Chipps, J. A., & Uys, L. (2014). Systematic review of health data quality management and best practices at community and district levels in LMIC. *Information Development*, *30*(2), 103-120.

Ngugi, P. N., Gesicho, M. B., & Babic, A. (2022, May). Impact of Electronic Medical Records Systems in Reporting HIV Health Data Indicators in Kenya. In *MIE* (pp. 234-238).

Noiki, I. A., Aladesanmi, L., Gupta, S., Mechael, P., Layer, E., Agarwal, S., & Gibson, D. (2025). Enablers and barriers to implementing a digital logistics system: A qualitative evaluation of OpenLMIS for Nigeria's vaccine supply chain. *medRxiv*, 2025-11.

Nwankwo, I. S. (2025). Cross-Border Health Data Transfers from Nigeria: Navigating the Legal and Ethical Landscape. In *International Transfers of Health Data: A Global Perspective* (pp. 195-218). Singapore: Springer Nature Singapore.

Odeny, B. M., Njoroge, A., Gloyd, S., Hughes, J. P., Wagenaar, B. H., Odhiambo, J., ... & Puttkammer, N. (2023). Development of novel composite data quality scores to evaluate facility-level data quality in electronic data in Kenya: a nationwide retrospective cohort study. *BMC Health Services Research*, *23*(1), 1139.

Ouedraogo, M., Kurji, J., Abebe, L., Labonté, R., Morankar, S., Bedru, K. H., ... & Kulkarni, M. A. (2019). A quality assessment of Health Management Information System (HMIS) data for maternal and child health in Jimma Zone, Ethiopia. *PloS one*, *14*(3), e0213600.

Oware, P. M., Omondi, G., Adipo, C., Adow, M., Wanyama, C., Odallo, D., ... & Were, F. (2025). Perceived accuracy and utilisation of DHIS2 data for health decision making and advocacy in Kenya: A Qualitative Study. *PLOS global public health*, *5*(8), e0004508.

Pandya, S. (2026). *From Transcription to Intelligence: Applying Large Language Models to Healthcare Enterprise Decision Support Systems*. Geh press.

Peters, M. D., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., ... & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI evidence synthesis*, *18*(10), 2119-2126.

Poppe, O., Sæbø, J. I., & Braa, J. (2021). WHO digital health packages for disseminating data standards and data use practices. *International Journal of Medical Informatics*, *149*, 104422.

Rogers, B. M. (2023). *Sub-national Health Management and Leadership Strengthening in Eastern and Southern Africa: Understanding the Enabling Environment*. Columbia University.

Rumisha, S. F., Lyimo, E. P., Mremi, I. R., Tungu, P. K., Mwingira, V. S., Mbata, D., ... & Mboera, L. E. (2020). Data quality of the routine health management information system at the primary healthcare facility and district levels in Tanzania. *BMC medical informatics and decision making*, *20*(1), 340.

Saini, V., Pal, D. K. D., & Reddy, S. G. (2022). Data quality assurance strategies in interoperable health systems. *J Artif Intell Res*, *2*(2), 322-358.

Setiawan, A. B., Nugroho, A. C., Ariansyah, K., Dunan, A., & Mudjiyanto, B. (2025). The use of big data for the health system in handling the COVID-19 pandemic in Indonesia. In *Digital Healthcare, Digital Transformation and Citizen Empowerment in Asia-Pacific and Europe for a Healthier Society* (pp. 31-44). Academic Press.

Shen, Y., Yu, J., Zhou, J., & Hu, G. (2024). 25 Years of Evolution and Hurdles in Electronic Health Records and Interoperability in Medical Research: A Comprehensive Review. *Journal of Medical Internet Research*, *27*(27).

Spangler, K. R., Levy, J. I., Fabian, M. P., Haley, B. M., Carnes, F., Patil, P., ... & Lane, K. J. (2023). Missing race and ethnicity data among COVID-19 cases in Massachusetts. *Journal of Racial and Ethnic Health Disparities*, *10*(4), 2071-2080.

Sukumar, S. R., Natarajan, R., & Ferrell, R. K. (2015). Quality of Big Data in health care. *International journal of health care quality assurance*, *28*(6), 621-634.

Syed, R., Eden, R., Makasi, T., Chukwudi, I., Mamudu, A., Kamalpour, M., ... & Myers, T. (2023). Digital health data quality issues: systematic review. *Journal of Medical Internet Research*, *25*, e42615.

Teklegiorgis, K., Tadesse, K., Terefe, W., & Mirutse, G. (2016). Level of data quality from Health Management Information Systems in a resource-limited setting and its associated factors, eastern

Ethiopia. *South African Journal of Information Management*, *18*(1), 1-8.

Thompson, H. S., Kinney, T. N., Leach, C., Sass, A., Washington, A., Dailey, R., ... & Khaldun, J. (2025). Use of Rapid Qualitative Analysis to Support Collaborative Synergy Within a Community Coalition for Health Equity in Detroit. *Health Equity*, *9*(1), heq-2024.

Tolera, A., Firdisa, D., Roba, H. S., Motuma, A., Kitesa, M., & Abaerei, A. A. (2024). Barriers to healthcare data quality and recommendations in public health facilities in Dire Dawa city administration, eastern Ethiopia: a qualitative study. *Frontiers in Digital Health*, *6*, 1261031.

Torab-Miandoab, A., Samad-Soltani, T., Jodati, A., & Rezaei-Hachesu, P. (2023). Interoperability of heterogeneous health information systems: a systematic literature review. *BMC medical informatics and decision making*, *23*(1), 18.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine*, *169*(7), 467-473.

Tucker, C. E., & Wang, Y. (2021). The role of delayed data in the COVID-19 pandemic. *Available at SSRN 3867021*.

Vasudevan, V., Gnanasekaran, A., Bansal, B., Lahariya, C., Parameswaran, G. G., & Zou, J. (2022). Assessment of COVID-19 data reporting on 100+ websites and apps in India. *PLOS Global Public Health*, *2*(4), e0000329.

Wagenaar, B. H., Hirschhorn, L. R., Henley, C., Gremu, A., Sindano, N., & Chilengi, R. (2017). Data-driven quality improvement in low-and middle-income country health systems: lessons from seven years of implementation experience across Mozambique, Rwanda, and Zambia. *BMC Health Services Research*, *17*(Suppl 3), 830.

Wang, J., Liu, Y., Li, P., Lin, Z., Sindakis, S., & Aggarwal, S. (2024). Overview of data quality: Examining the dimensions, antecedents, and impacts of data quality. *Journal of the Knowledge Economy*, *15*(1), 1159-1178.

Wiley Jr, K. K. (2022). *Electronic Health Record (EHR) Data Quality and Type 2 Diabetes Mellitus Care*. Indiana University-Purdue University Indianapolis.

World Health Organization. (2025). *Advancing the responsible use of digital technologies in global health: A report of the WHO Science Council*. World Health Organization.

Zhang, G., Zhao, Z., Zhang, X., Wu, X., Zheng, Y., Feng, L., & Huang, Z. (2024c). Comprehensive Multi-Source remote sensing data integration for enhanced mineralization alteration extraction and geological structure interpretation in the Lala region of Sichuan Province. *Ore Geology Reviews*, *168*, 106032.

Zhang, Y., Callaghan-Koru, J. A., & Koru, G. (2024a). The challenges and opportunities of continuous data quality improvement for healthcare administration data. *JAMIA open*, *7*(3), ooae058.

Zhang, Y., Gao, J., Tan, Z., Zhou, L., Ding, K., Zhou, M., ... & Wang, D. (2024b). Data-centric foundation models in computational healthcare: A survey. *ACM Computing Surveys*.