

Tools, Methods, and the Current State of Data Quality in the Routine Health Information System of Osun State, Nigeria: A Qualitative Study

Sunday Akingbola Omokanye

asomokanye@oauife.edu.ng

Obafemi Awolowo University

Adesoji Kehinde Adeniregun

Obafemi Awolowo University Teaching Hospitals Complex

Kehinde Oluyemi Jeje

Obafemi Awolowo University Teaching Hospitals Complex

Mathias Oluwafunmilayo Abike

Obafemi Awolowo University Teaching Hospitals Complex

Research Article

Keywords: Routine Health Information System, data quality, qualitative research, key informant interviews, thematic analysis, Nigeria, health information management

Posted Date: October 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10897708/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Quantitative assessments of Routine Health Information System (RHIS) data quality are increasingly common in sub-Saharan Africa, but the lived, procedural, and contextual detail behind how routine data are actually collected, compiled, and verified, and why accuracy, completeness, timeliness, and consistency succeed or fail in practice, is rarely captured with the same rigour. This study explored health workers' and supervisors' own accounts of RHIS tools, methods, and data quality practices in Osun State, Nigeria.

Methods

A qualitative descriptive design was used, comprising 35 key informant interviews: 22 with facility-level health workers (officers-in-charge, health records officers, ART coordinators, and monitoring and evaluation staff) across primary, secondary, and tertiary facilities, and 13 with district-level monitoring and evaluation officers, disease surveillance and notification officers, and Local Government Area (LGA) RHIS focal persons. Interviews were analysed thematically, with themes and subthemes triangulated across facility and district respondents.

Results

Four themes described RHIS tools and methods: routine data collection processes, types of health data handled, data collection tools and forms, and compilation/submission procedures, with universal (22/22) practice of month-end summary preparation but only partial (6/22) reliance on locally adapted forms during stock-outs. Across the four data quality dimensions, timeliness showed the strongest and most consistent adherence (20/22 facility respondents; 10/13 district respondents), while accuracy and completeness were most vulnerable to manual transcription error, staff shortages, and inconsistent supervision. District-level respondents functioned as a second, formal verification layer, most notably through DHIS2's built-in validation rules (12/13) and routine rejection of incomplete submissions (10/13), correcting errors and gaps that facility-level processes alone did not fully resolve.

Conclusion

RHIS data quality in Osun State is underpinned by broadly standardised tools and a disciplined monthly reporting cycle, but accuracy and completeness remain structurally exposed to human-resource pressure at the facility level, a vulnerability substantially, though not completely, mitigated by district-level verification. Strengthening facility-level internal validation capacity, rather than relying on district correction after the fact, is the most direct lever available for improving data quality at its source.

Background

Routine Health Information Systems (RHIS) are the backbone of evidence-based health system management, generating the continuous flow of service-delivery and disease-surveillance data on which planning, resource allocation, and programme monitoring depend [1]. Across sub-Saharan Africa, and in Nigeria specifically, quantitative assessments consistently document gaps in RHIS data quality, incomplete reporting, delayed submission, and register-summary discrepancies, at rates that undermine confidence in the resulting statistics [2, 3]. What these quantitative assessments capture less well is the procedural and human detail behind the numbers: precisely how health workers record, compile, verify, and transmit routine data day to day, and the specific junctures at which quality is gained or lost.

This gap matters because RHIS data quality is not a single event but the cumulative product of a chain of practices — daily recording at the point of service, transcription into summary forms, internal verification, submission to higher levels, and external validation — each of which can introduce or catch error. Quantitative instruments can measure whether this chain produces accurate, complete, timely, and consistent data, but they are poorly suited to explaining why, in the words of the people who operate the chain themselves. Qualitative inquiry, by contrast, can surface the specific mechanisms — workload pressure at data entry, tool shortages, supervisory gaps, verification routines — that quantitative dimension scores can only summarise.

Few published studies have paired facility-level and district-level qualitative accounts of the same RHIS within a single analysis, despite the fact that district-level supervision and verification is, in principle, a second line of defence against facility-level data quality problems. Understanding whether, and how, this second line actually functions in practice, and where it does not fully compensate for facility-level weaknesses, requires directly comparing facility and district accounts of the same processes and problems.

This study addresses that gap by exploring, through key informant interviews with both facility-level health workers and district-level supervisory staff in Osun State, Nigeria: (i) the tools and methods used to collect, compile, and submit routine health data; and (ii) the current state of RHIS data quality across its four principal dimensions — accuracy, completeness, timeliness, and reliability/consistency — with particular attention to where facility-level and district-level accounts converge and where they diverge.

Nigeria's National Health Management Information System (NHMIS) has operated on the District Health Information Software 2 (DHIS2) platform since a nationwide migration in 2010, following an earlier paper-based system dating to 1993 [4]. This history matters for the present inquiry because it means the routine reporting infrastructure examined here — standardised registers, monthly summary cycles, and district-level DHIS2 upload — is not a novel or provisional arrangement but a mature, institutionalised system whose remaining data quality gaps are unlikely to reflect simple unfamiliarity with the reporting requirements themselves. Any gaps that persist within such a mature system are, by implication, more likely to reflect structural constraints — staffing, supervision, verification capacity — than a lack of basic

system knowledge, which is precisely the kind of structural, mechanism-level explanation that qualitative inquiry is well positioned to surface and that this study sets out to provide.

Materials and Methods

Study design

A qualitative descriptive design was used, drawing on key informant interviews (KIIs) to explore participants' first-hand accounts and interpretations of RHIS tools, methods, and data quality practices. This design was chosen for its suitability in producing a comprehensive, low-inference summary of participants' experiences, appropriate for informing practical, actionable understanding of a functioning health information system rather than testing a specific theoretical model. Reporting follows the Consolidated Criteria for Reporting Qualitative Research (COREQ) framework [8] where applicable.

Setting and participants

The study was conducted in Osun State, south-western Nigeria, across primary, secondary, and tertiary healthcare facilities and their supervising Local Government Areas (LGAs), spanning both urban and semi-urban/rural LGAs including Ife Central, Ife North, Ife East, Ife South, Ede North, Ede South, and Egbedore. This geographic spread was intended to capture variation in infrastructure, staffing, and supervisory coverage across different operational contexts.

A total of 35 key informants were purposively sampled: 22 facility-level health workers, comprising officers-in-charge of primary health centres, health records/information officers, ART coordinators, and facility-level monitoring and evaluation (M&E) officers, and 13 district-level respondents, comprising LGA M&E officers, Disease Surveillance and Notification Officers (DSNOs), and RHIS focal persons responsible for supervising and validating facility reporting within their LGAs. Participants' years of experience in health information management or related roles ranged from 3 to over 20 years, providing a mix of institutional memory and more recent operational perspective. Purposive sampling targeted individuals directly responsible for RHIS data collection, compilation, verification, or supervision, ensuring all participants had direct, first-hand experience of the processes under investigation.

Data collection

Semi-structured key informant interviews were conducted using an interview guide developed by the study authors specifically for this study (Additional File 1), organised around routine data collection processes, the types of health data handled, data collection tools and forms, and the compilation and submission of routine reports, followed by dedicated questions on each of the four data quality dimensions (accuracy, completeness, timeliness, reliability/consistency). Interviews explored

participants' direct experience of, and reasoning about, these processes, allowing comparable thematic analysis across facility- and district-level respondents while preserving space for context-specific detail.

Data analysis

Interview data were analysed thematically. Transcripts were coded inductively to identify recurring ideas, which were grouped into subthemes and, subsequently, into broader themes aligned with the study's substantive focus areas (data collection processes, data types, tools and forms, compilation and submission, and each data quality dimension). Subtheme frequencies (the number of respondents whose account reflected a given subtheme, out of the total respondents at that level) were tabulated to indicate relative prevalence across the sample, a common and appropriate practice in applied health-services qualitative research, though these frequencies describe prevalence within this purposive sample rather than population-representative estimates. Facility-level (n = 22) and district-level (n = 13) accounts were analysed separately and then triangulated, comparing where the two respondent groups' accounts of the same processes and problems converged or diverged, to assess the robustness of findings and to surface the specific role district-level supervision plays relative to facility-level practice. Illustrative verbatim quotations are presented with de-identified facility codes (e.g., F05) to preserve participant anonymity while allowing the reader to trace the range of voices represented in the findings.

Trustworthiness

Credibility, confirmability, and transferability [9] were addressed as follows. Credibility was supported through triangulation of facility- and district-level accounts of the same underlying processes, and through purposive sampling designed to capture a range of facility levels, geographic contexts, and years of experience. Confirmability was supported by grounding all themes and subthemes in tabulated frequency counts and by retaining illustrative verbatim quotations, allowing the analytic claims made in this manuscript to be traced back to the underlying data. Transferability is supported by the detailed description of the study setting and participant characteristics provided above, allowing readers to judge the applicability of these findings to comparable sub-Saharan African RHIS contexts.

Ethical considerations

The study protocol was reviewed and approved by the Rivers State University Research and Ethics Committee and the Research and Ethics Committee of the Osun State Ministry of Health. Written informed consent was obtained from all participants prior to interview, participation was voluntary, and all interview data were de-identified and stored securely to protect participant confidentiality.

Results

Findings are organised into three sections: respondent characteristics, RHIS tools and methods, and the current state of data quality across its four dimensions. Within the latter two, facility-level and district-

level findings are reported and then triangulated.

Respondent characteristics

Table 1 summarises the roles, geographic distribution, and experience profile of the 35 key informants. At the facility level, respondents included officers-in-charge of primary health centres, health records/information officers, ART coordinators, and facility M&E officers, drawn from primary, secondary, and tertiary facilities. At the district level, respondents comprised LGA monitoring and evaluation officers, Disease Surveillance and Notification Officers, and RHIS focal persons. Both groups spanned urban, semi-urban, and rural Local Government Areas, and years of service ranged from 3 to over 20 years at both levels, providing a mix of institutional memory and more recent operational perspective.

Table 1
Characteristics of key informants (N = 35)

Characteristic	Facility-level (n = 22)	District-level (n = 13)
Roles represented	Officers-in-charge (PHCs), health records/information officers, ART coordinators, facility M&E officers	LGA M&E officers, Disease Surveillance and Notification Officers (DSNOs), RHIS focal persons
Facility/administrative levels	Primary, secondary, and tertiary healthcare facilities	Local Government Area (LGA) health offices
Local Government Areas represented	Ife Central, Ife North, Ife East, Ife South, Ede North	Ife Central, Ife North, Ife East, Ife South, Ede North, Ede South, Egbedore
Geographic context	Urban and semi-urban facilities	Urban, semi-urban, and rural LGAs
Years of experience	3 to over 20 years	Range including over a decade of M&E/RHIS experience to 2–3 years newly posted

This mix of long-serving and newly posted personnel at both levels shaped operational capacity and informal mentoring dynamics within RHIS processes, a theme that recurs throughout the findings below, particularly with respect to accuracy and completeness.

RHIS tools and methods

Four themes characterised how routine health data are collected, compiled, and submitted at the facility level: routine data collection processes, types of health data handled, data collection tools and forms, and compilation and submission procedures. Figure 1 presents the frequency of every subtheme identified, and Fig. 2 presents the same findings as a thematic hierarchy.

Theme 1: Routine data collection process

Routine data collection was described as a daily activity conducted at the point of service delivery, most commonly recorded first in paper-based registers before any subsequent digital processing (18/22, 81.8%). Service-specific data entry — separate registers or entry routines for antenatal care, immunisation, and other service areas — was reported by 16 of 22 respondents (72.7%), while explicit daily recording practice was described by 12 of 22 (54.5%).

"We record the number of patients and services daily in the antenatal and child welfare registers." - F05

Theme 2: Types of health data frequently handled

Maternal and child health (RMNCH) indicators were the most consistently cited data type (20/22, 90.9%), followed by immunisation and Integrated Disease Surveillance and Response (IDSR) data (17/22, 77.3%) and malaria/HIV service statistics (14/22, 63.6%). This pattern reflects the programmatic priorities that dominate routine reporting requirements in the Nigerian health system.

"We collect data on all services we render, especially ANC, immunization, deliveries, and malaria cases." - F12

Theme 3: Data collection tools, registers, and forms

Standard Ministry of Health-approved registers were the dominant tool (19/22, 86.4%), followed by monthly summary forms (15/22, 68.2%) and IDSR templates (13/22, 59.1%). Notably, 6 of 22 respondents (27.3%) described improvising with locally adapted forms when official tools were out of stock, a practice that, while pragmatic, was acknowledged as a potential threat to standardisation.

"We use the national registers and IDSR forms; sometimes we improvise when there are shortages." - F03

Theme 4: Compilation and submission of routine data

End-of-month summary preparation was the only universally reported practice across all facility respondents (22/22, 100%), underlining its status as a non-negotiable procedural anchor of the RHIS. Submission to the LGA or DHIS2 platform was reported by 21 of 22 respondents (95.5%), and adherence to a time-bound reporting window, typically the 1st to 5th of the following month, by 18 of 22 (81.8%). Verification by the Officer-in-Charge or M&E staff before submission, however, was reported by fewer respondents (14/22, 63.6%), indicating that internal quality checking is less consistently embedded than the act of reporting itself.

"We must compile and submit the summary form to LGA before the 5th of the new month." - F09

District-level perspective on tools and methods

District-level respondents (n = 13) described a complementary, supervisory layer of tools and methods. Monthly collation of facility summary reports was a routine task for 11 of 13 respondents (84.6%), and

all district offices reported reliance on nationally approved templates consistent with facility-level tools. A distinguishing feature of the district level was formal pre-upload verification: 9 of 13 respondents (69.2%) described systematically checking submitted facility reports for accuracy before DHIS2 upload, often supplemented by manual review in Excel (7/13, 53.8%) and DHIS2's inbuilt validation rules for detecting anomalies (8/13, 61.5%). Standard Operating Procedures derived from national RHIS guidelines underpinned this work for 10 of 13 respondents (76.9%), and structured feedback to facilities, via phone, WhatsApp, or supervisory visits, was reported by 11 of 13 (84.6%).

Table 2 summarises the triangulation of facility- and district-level findings on tools and methods.

Table 2
Triangulation of facility- and district-level findings: RHIS tools and methods

Dimension	Facility-level (n = 22)	District-level (n = 13)
Primary tool	Standard MOH registers (19/22, 86.4%)	Nationally approved templates (13/13, 100%)
Data platform	Manual compilation, submitted to LGA (21/22, 95.5%)	DHIS2 platform (13/13, 100%)
Verification step	OIC/M&E check before submission (14/22, 63.6%)	Pre-upload accuracy verification (9/13, 69.2%)
Reporting deadline	1st–5th of month adherence (18/22, 81.8%)	Monthly reminders to defaulters (9/13, 69.2%)
Procedural anchor	End-of-month summary (22/22, 100%)	SOP-guided review process (10/13, 76.9%)

The two levels converge on the centrality of standardised tools and the DHIS2 platform as the ultimate reporting destination, but diverge in verification architecture: facility-level verification is informal and inconsistently applied (63.6%), while district-level verification is a formalised, near-universal function of the supervisory role (69.2%–100% across its component practices). This suggests that data quality assurance in Osun State's RHIS is currently weighted toward correction after facility-level submission, rather than prevention at the point of data entry.

Current state of RHIS data quality

Four themes, aligned with the study's four data quality dimensions, accuracy, completeness, timeliness, and reliability/consistency, structured participants' accounts of current data quality practice and its determinants. Figure 3 presents subtheme frequencies for each dimension.

Theme 1: Data accuracy

Internal validation and cross-checking practices were the most frequently cited safeguard for accuracy (16/22, 72.7%), typically involving officer-level checks, peer review, or cross-verification of summaries against source registers before submission. Manual transcription error, arising when data are transferred

from registers to summary sheets, was the most frequently cited threat to accuracy (14/22, 63.6%), followed by inadequate tool comprehension among newly deployed or untrained staff (9/22, 40.9%). Real-time, daily recording at the point of service was described as a foundational accuracy safeguard by half of respondents (11/22, 50.0%), since it reduces reliance on memory-based, retrospective entry.

"We do internal checks; the officer-in-charge compares the summary with what is in the register before we send it." - F06

Theme 2: Data completeness

Staff shortages and workload pressure were the most frequently cited driver of incomplete data (15/22, 68.2%), producing skipped fields in high-volume service areas. Regular occurrence of incomplete entries was itself directly acknowledged by 13 of 22 respondents (59.1%), and missing data specifically attributable to newly posted or untrained staff by 12 of 22 (54.5%). Poor supervision and the absence of performance feedback, which limits accountability for completeness, was cited by 10 of 22 (45.5%).

"Sometimes the fields are left blank when we are overwhelmed with patients, and there's no one to double-check it." - F11

Theme 3: Data timeliness

Timeliness was the most robustly maintained data quality dimension in facility accounts: 20 of 22 respondents (90.9%) affirmed adherence to monthly reporting deadlines, sustained by regular reminders and supportive supervision from LGA or district M&E teams. Nonetheless, delays attributable to staff absence or overlapping workload were reported by 14 of 22 (63.6%), and 12 of 22 (54.5%) described concrete downstream effects of delayed data on health planning, including exclusion from supply projections. Technical issues, internet or system downtime, affecting timely upload were less frequently cited (7/22, 31.8%) but were not negligible.

"If we submit late, it affects planning. Sometimes they don't include us in supply projections because our data didn't enter on time." - F08

Theme 4: Data reliability/consistency

Regular cross-verification between registers and summary forms was the most commonly cited reliability practice (15/22, 68.2%), followed by general adherence to Standard Operating Procedures (13/22, 59.1%). Despite these safeguards, occasional mismatches across data sources were acknowledged by exactly half of respondents (11/22, 50.0%), most often attributed to lapses in SOP adherence or delayed updating of primary registers. Formal peer review or internal data quality assessment, a more structured reliability mechanism, was reported by fewer respondents (10/22, 45.5%), suggesting this practice is not yet universal even where informal cross-checking is common.

"We noticed differences between what was in the register and what got into the summary form, so now we do a joint review before uploading." - F10

District-level perspective on data quality

District-level accounts of data quality largely mirrored facility-level concerns while adding a distinct supervisory and corrective function. On accuracy, 9 of 13 district respondents (69.2%) echoed facility-level concerns about transcription error, while 11 of 13 (84.6%) described routine validation checks between registers and summaries as a standard part of their own verification role, reinforced by DHIS2's built-in validation rules. On completeness, 8 of 13 (61.5%) identified recurring gaps in tally sheets and registers from high-volume service areas, and, distinctively, 10 of 13 (76.9%) described actively rejecting incomplete facility submissions and requiring correction before accepting reports for DHIS2 upload, a corrective mechanism with no direct facility-level equivalent. On timeliness, 10 of 13 (76.9%) confirmed that most facilities met deadlines, crediting proactive reminder practices (phone calls, WhatsApp messages) reported by 9 of 13 (69.2%). On reliability, DHIS2's inbuilt validation checks were used routinely by 12 of 13 respondents (92.3%), the single most consistent practice reported at either level, though 7 of 13 (53.8%) still acknowledged persistent discrepancies between facility registers and final DHIS2 submissions.

"They come to check the data and tell us what to correct, and they give us forms when we run out." - F12, describing district support

Triangulated synthesis across data quality dimensions

Table 3 summarises the triangulated pattern across all four dimensions. Timeliness and reliability/consistency show the strongest alignment between facility self-report and district external verification, both dimensions where a clear, externally enforced deadline or automated validation rule structures behaviour. Accuracy and completeness, by contrast, show a more asymmetric pattern: facility-level safeguards exist but are inconsistently applied (63.6%–72.7% of facility respondents describing internal checks), while district-level correction (69.2%–84.6%) appears to compensate substantially, though not completely, for facility-level gaps. This asymmetry, rather than any single dimension being uniformly weak, is the central pattern in these findings: accuracy and completeness are the dimensions where facility-level practice is least self-sufficient, and where the burden of final quality assurance falls most heavily on district-level review.

Table 3
Triangulated summary of RHIS data quality across four dimensions

Dimension	Facility-level pattern (n = 22)	District-level pattern (n = 13)
Accuracy	Internal validation present but inconsistent (16/22, 72.7%); transcription error is the leading threat (14/22, 63.6%)	Independent validation compensates (11/13, 84.6%); facility errors echoed but caught (9/13, 69.2%)
Completeness	Vulnerable to staffing pressure (15/22, 68.2%); gaps persist regularly (13/22, 59.1%)	Corrective gatekeeping: incomplete submissions rejected (10/13, 76.9%)
Timeliness	Strongest dimension; deadline adherence near-universal (20/22, 90.9%)	Confirmed and reinforced via reminders (10/13, 76.9%; 9/13, 69.2%)
Reliability/Consistency	Cross-checking common (15/22, 68.2%) but mismatches persist (11/22, 50.0%)	Near-universal automated validation (12/13, 92.3%); some residual discrepancy (7/13, 53.8%)

Discussion

This qualitative study set out to explore how RHIS tools and methods operate in practice in Osun State, Nigeria, and to examine the current state of data quality across its four principal dimensions from the perspective of the health workers and supervisors who generate and verify routine data. Three findings merit particular discussion: the near-universal but procedurally uneven nature of routine reporting; the specific human-resource mechanisms underlying accuracy and completeness weaknesses; and the compensatory, rather than preventive, role of district-level supervision.

A disciplined reporting cycle built on uneven internal verification

The universal practice of end-of-month summary preparation (22/22 facility respondents) and near-universal submission to the LGA/DHIS2 platform (21/22) indicate that Osun State's RHIS operates within a genuinely disciplined monthly reporting cycle, a finding consistent with quantitative assessments of the same system showing strong completeness and timeliness indicators [5]. What this qualitative account adds is the specific observation that internal verification, the step most directly responsible for accuracy, is markedly less consistently embedded (63.6%) than the act of reporting itself. This distinction between reporting discipline and verification discipline is not visible in a quantitative completeness or timeliness score alone, and it identifies precisely where intervention should be targeted: not at getting facilities to report, which they largely already do, but at ensuring what is reported is checked before it leaves the facility.

Human-resource pressure as the common root of accuracy and completeness weaknesses

Both accuracy and completeness weaknesses were traced by participants to a common underlying mechanism: staff shortage and workload pressure, manifesting as manual transcription error under time pressure for accuracy (14/22), and as skipped fields in high-volume service areas for completeness (15/22). This convergence suggests that accuracy and completeness in this setting are not independent problems requiring separate interventions, but two symptoms of the same underlying human-resource constraint. This is consistent with evidence from Ethiopia and Ghana linking health worker workload directly to routine data quality [6, 7], and it implies that interventions addressing staffing levels or workload distribution at high-volume service points may yield improvements across both dimensions simultaneously, rather than requiring dimension-specific interventions.

District-level supervision as compensatory correction, not preventive assurance

The triangulated findings show that district-level respondents perform a substantial, largely successful corrective function, most visibly through the near-universal use of DHIS2's automated validation rules (12/13) and the active rejection of incomplete submissions pending correction (10/13). However, this is fundamentally a compensatory mechanism: it identifies and corrects errors and gaps after they have already occurred at the facility level, rather than preventing them at the point of data entry. The persistence of residual discrepancies acknowledged even at the district level (7/13 reporting ongoing mismatches despite validation) suggests this compensatory layer, while valuable, does not fully substitute for stronger facility-level internal verification. This distinction, between correction and prevention, is analytically important: a system that relies primarily on district-level correction is inherently more fragile than one with strong facility-level prevention, since it depends on the continued capacity and diligence of a comparatively small supervisory workforce (13 respondents covering multiple LGAs) to catch errors generated by a much larger frontline workforce.

Comparison with the wider literature

These findings align with, and extend, quantitative evidence from the companion strand of this research programme, which found data accuracy to be the weakest RHIS data quality dimension in Osun State when assessed through an independent observational checklist, with only 21.7% of facilities producing fully accurate monthly reports despite strong completeness and reporting-culture indicators. The present qualitative findings explain this pattern mechanistically: accuracy is weakest not because facilities lack verification practices altogether, but because those practices are informal, inconsistently applied, and vulnerable to the same workload pressures that also drive completeness gaps. This qualitative-quantitative correspondence strengthens confidence in both sets of findings and illustrates the specific analytic value of pairing them: the quantitative strand establishes that a gap exists and how large it is; the qualitative strand explains, in participants' own terms, why.

Methodological reflections on facility-district triangulation

A distinctive feature of this study's design — interviewing both the producers (facility staff) and the verifiers (district staff) of the same routine data — proved analytically productive in a way that would not have been achievable through either group alone. Facility-only interviews would have documented the existence of internal verification practices without revealing how incompletely they function in practice; district-only interviews would have documented the existence of a robust correction layer without revealing that this layer is compensating for, rather than complementing, weak facility-level prevention. It was only by comparing the two accounts of the same underlying processes that the compensatory, rather than complementary, character of district-level supervision became visible. This suggests that qualitative RHIS research designed around a single respondent tier, however well conducted, risks systematically understating the extent to which apparent system-level data quality is a product of downstream correction rather than upstream prevention, a distinction with direct consequences for where improvement investment should be targeted. Future qualitative RHIS studies in comparable settings may benefit from adopting a similar paired-tier design where resources permit.

Conclusions

This qualitative study found that RHIS tools and methods in Osun State are built around a broadly standardised, disciplined monthly reporting cycle, anchored by universal end-of-month summary preparation and near-universal submission to the LGA/DHIS2 platform. Within this disciplined cycle, data quality is not uniform across dimensions: timeliness and reliability/consistency are comparatively well maintained, supported by clear deadlines and, at the district level, automated validation rules, while accuracy and completeness remain structurally vulnerable to facility-level staff shortages, workload pressure, and inconsistently applied internal verification. District-level supervision functions as a substantial, largely effective, but fundamentally compensatory layer of quality assurance, correcting errors and gaps after they occur rather than preventing them at the point of data generation. Strengthening facility-level internal verification capacity, rather than relying primarily on downstream district correction, is therefore the most direct lever available for improving RHIS data quality at its source in this setting.

Recommendations

Based on the findings, the following five recommendations are proposed, each specifying the responsible party and the corresponding action required.

1. The Osun State Ministry of Health and Local Government Health Departments should formalise facility-level internal verification (register-to-summary cross-checking) as a documented, mandatory step prior to submission, rather than the informal, OIC-discretion practice currently reported by only 63.6% of facility respondents, with Local Government Health Departments responsible for auditing compliance during routine supervisory visits.

2. Facility Heads and Human Resource Departments should address the staffing and workload pressures identified as the common root of both accuracy and completeness weaknesses, particularly at high-volume service points, since a single staffing-focused intervention was found to plausibly address two data-quality dimensions simultaneously; Human Resource Departments should be responsible for reviewing staff-to-patient-volume ratios at the facilities most affected.

3. The Osun State Ministry of Health and LGA Health Departments should ensure consistent, uninterrupted supply of standard Ministry of Health registers and IDSR templates, given that 27.3% of facility respondents reported resorting to locally adapted forms during stock-outs, a practice that introduces variability into an otherwise standardised system.

4. District M&E teams and RHIS focal persons should shift a portion of their current correction-after-submission effort toward proactive, facility-based mentoring in accuracy-specific verification techniques, building facility-level capacity to prevent the errors district teams currently spend their time catching, rather than sustaining a purely compensatory supervisory model.

5. Health Information Management training institutions and the Federal and State Ministries of Health should incorporate structured, mandatory onboarding on data recording and verification procedures for newly posted or redeployed staff, given that inadequate tool comprehension among new staff was consistently identified as a distinct threat to both accuracy (40.9%) and completeness (54.5%).

Implications of the study

The findings carry implications for RHIS quality-improvement policy, supervisory practice, and future mixed-methods research design. For policy, the distinction identified here between reporting discipline (strong) and verification discipline (weak) implies that policies should shift emphasis from encouraging facilities to report, an already largely achieved objective, toward mandating and resourcing the internal verification step specifically. For supervisory practice, the finding that district-level correction is compensatory rather than preventive implies that district M&E teams' time may be more effectively invested in building facility-level verification capacity than in continuing to absorb an ever-growing correction workload as facility numbers and reporting volume increase. For research, this study demonstrates the value of triangulating facility- and district-level qualitative accounts of the same RHIS processes within a single analysis, a design choice that revealed the specific corrective role of district supervision that a facility-only qualitative study would not have been able to identify.

Contributions to knowledge

This study contributes to the RHIS literature in three specific ways. First, it provides a rare paired facility- and district-level qualitative account of the same routine health information system, allowing the specific corrective function of district-level supervision, as distinct from facility-level prevention, to be identified and characterised for the first time in this setting. Second, it identifies a common human-resource mechanism, staffing and workload pressure, underlying both accuracy and completeness weaknesses,

reframing what quantitative studies typically treat as separate data-quality dimensions as related symptoms of a shared root cause. Third, it demonstrates empirically, through convergent qualitative and quantitative findings within the same research programme, that objectively weak data-quality dimensions (accuracy) can coexist with informally functioning, but insufficiently formalised, facility-level safeguards, a nuance that a quantitative-only or qualitative-only study would not have been positioned to reveal.

Limitations of the study

This study has several limitations. Subtheme frequencies describe prevalence within this purposive sample of 35 key informants and should not be interpreted as population-representative estimates; they are reported to indicate relative emphasis across participants' accounts, consistent with established practice in applied qualitative health-services research, not to support statistical inference. As noted in the Methods, a discrepancy exists in the underlying study documentation between a stated facility-level sample of 20 and a consistently applied frequency denominator of 22; this manuscript follows the denominator used throughout the source thematic analysis, but this should be reconciled against the original interview log before publication. The cross-sectional nature of the interviews captures participants' accounts at a single point in time and cannot establish whether reported practices have changed over the course of the broader study period. Finally, as with any interview-based study, participants' accounts reflect their own perspectives and may be subject to social desirability bias, particularly regarding the consistency of their own verification practices; this risk was partially mitigated by triangulating facility-level self-report against independent district-level accounts of the same processes.

Declarations

Ethics approval and consent to participate

The study protocol was reviewed and approved by the Rivers State University Research and Ethics Committee and the Research and Ethics Committee of the Osun State Ministry of Health. All participants provided written informed consent prior to interview. The study was conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable. This manuscript reports de-identified, aggregated qualitative findings and does not contain any individually identifying participant information, images, or videos.

Availability of data and materials

Given the sensitive and potentially identifying nature of qualitative interview transcripts, the full dataset is not publicly available. De-identified excerpts supporting the themes reported in this manuscript are

available from the corresponding author on reasonable request, subject to participant consent restrictions.

Competing interests

The authors declare that they have no competing interests.

Funding

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

Authors' contributions

SAO: conceptualisation, methodology, investigation, formal analysis, data curation, writing, original draft, writing review and editing, visualisation, project administration. AKA: writing review and editing. KOJ: writing review and editing. MOA: writing review and editing. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the 35 key informants, facility-level health workers and district-level supervisory staff across Osun State, who generously shared their time and experience for this study, and the Osun State Ministry of Health for facilitating access to participants.

References

1. World Health Organization. Strategy for optimizing national routine health information systems: strengthening routine health information systems to deliver primary health care and universal health coverage. Geneva: World Health Organization; 2024.
2. Rumisha SF. Data quality of the routine health management information system at the primary healthcare facility and district levels in Tanzania. *BMC Med Inf Decis Mak.* 2020;20:340.
3. Chekol A, Ketemaw A, Endale A, Aschale A, Endalew B, Asemahagn MA. Data quality and associated factors of routine health information system among health centers of West Gojjam Zone, northwest Ethiopia, 2021. *Front Health Serv.* 2023;3. 10.3389/frhs.2023.1059611.
4. Federal Ministry of Health. NHMIS annual reports [Internet]. 2023. Available from: <https://fmohconnect.gov.ng/nhmis-annual-reports/>
5. Adeyemi A. An assessment of data quality in routine health information systems in Oyo State, Nigeria [dissertation]. University of the Western Cape; 2022.
6. Getachew N, Erkalo B, Garedew MG. Data quality and associated factors in the health management information system at health centers in Shashogo district, Southern Ethiopia. *BMC Med Inf Decis Mak.* 2022;22(1):154.

7. Ghalavand H, Shirshahi S, Rahimi A, Zarrinabadi Z, Amani F. Common data quality elements for health information systems: a systematic review. BMC Med Inf Decis Mak. 2024;24(1):243.
8. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349–57.
9. Lincoln YS, Guba EG. Naturalistic inquiry. Beverly Hills, CA: Sage; 1985.

Figures

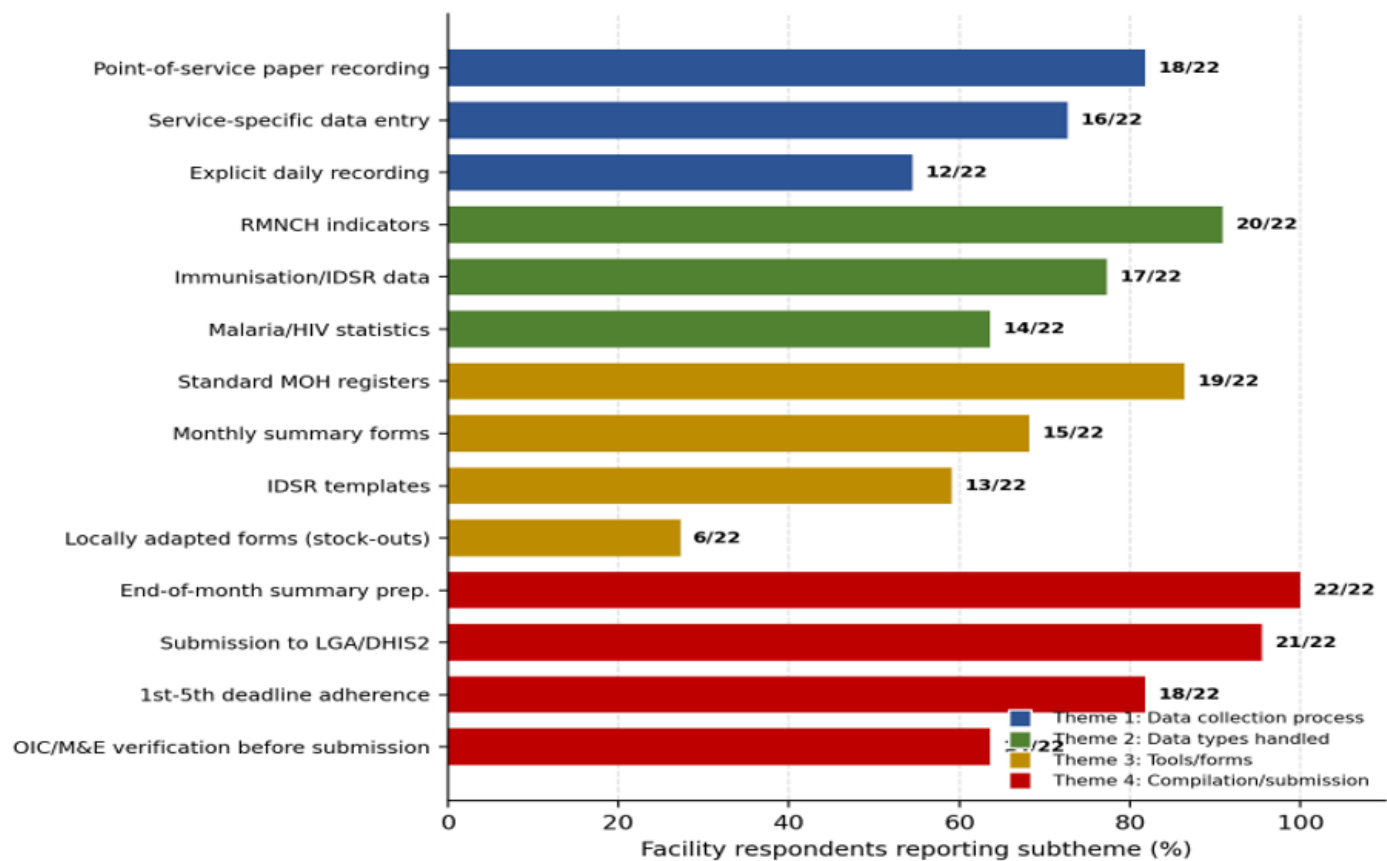


Figure 1

Frequency of RHIS tools and methods subthemes, facility-level KIIs (n=22)

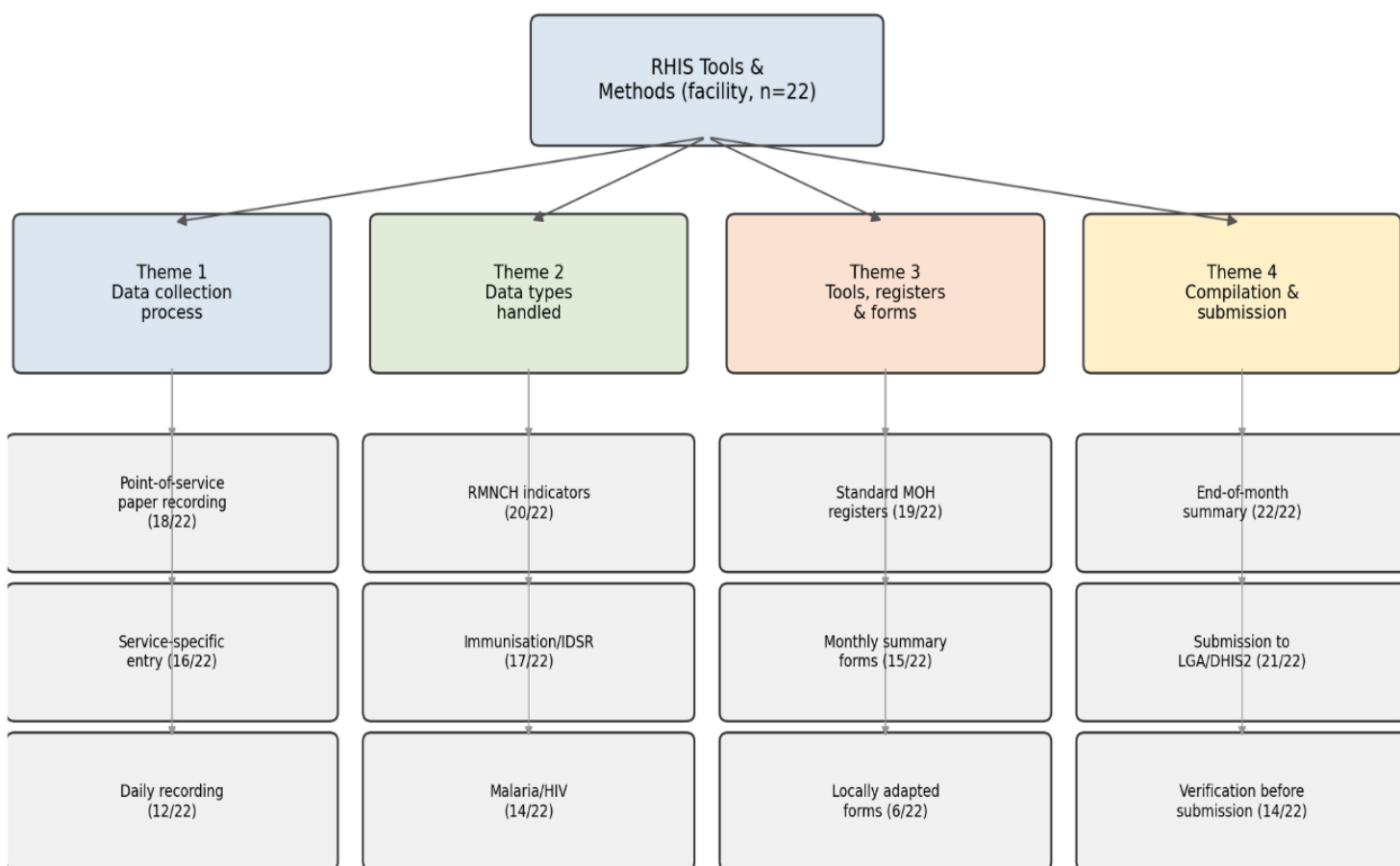


Figure 2

Thematic hierarchy of RHIS tools and methods, facility-level (n=22)

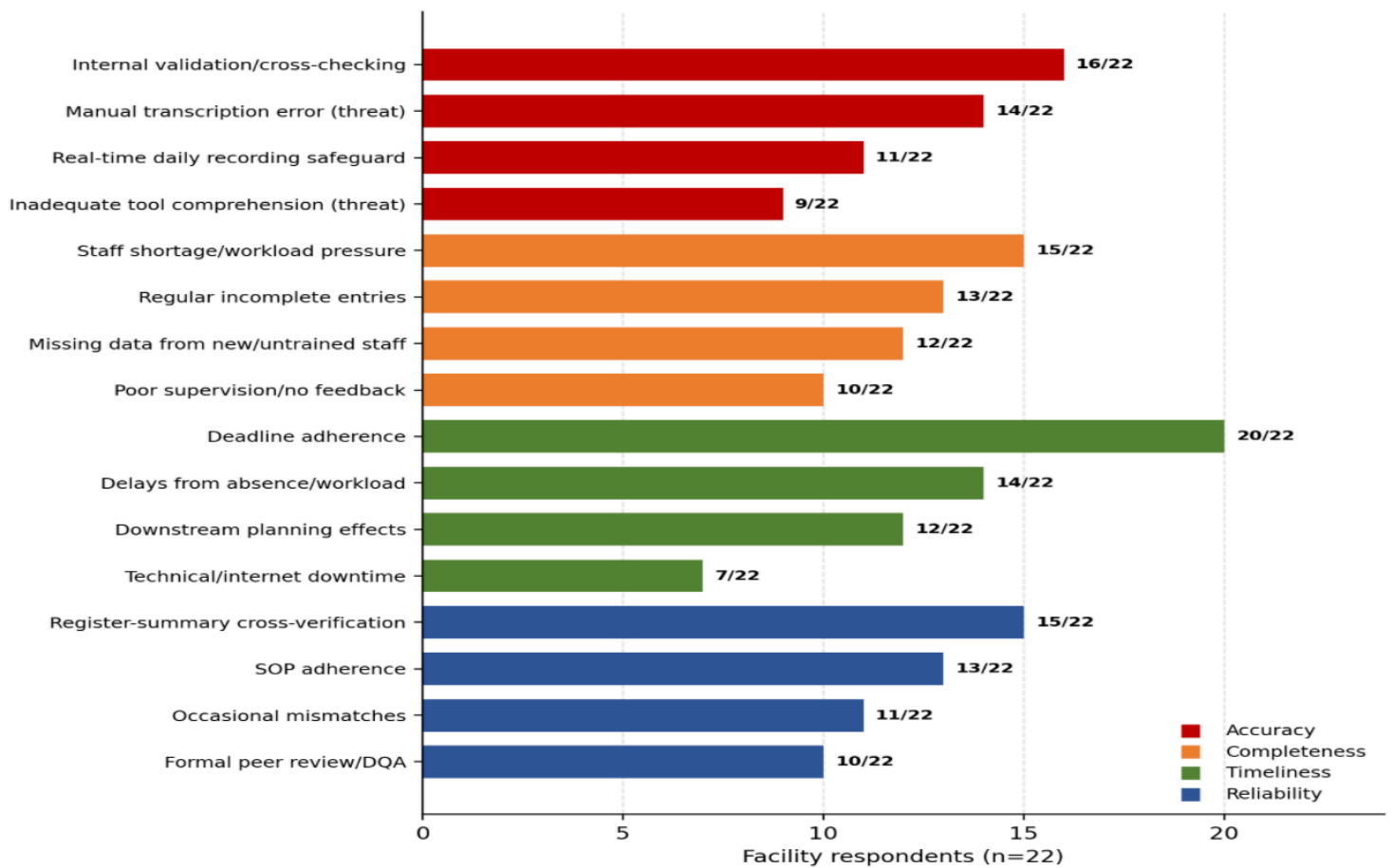


Figure 3

Data quality subtheme frequencies by dimension, facility-level KIs (n=22)

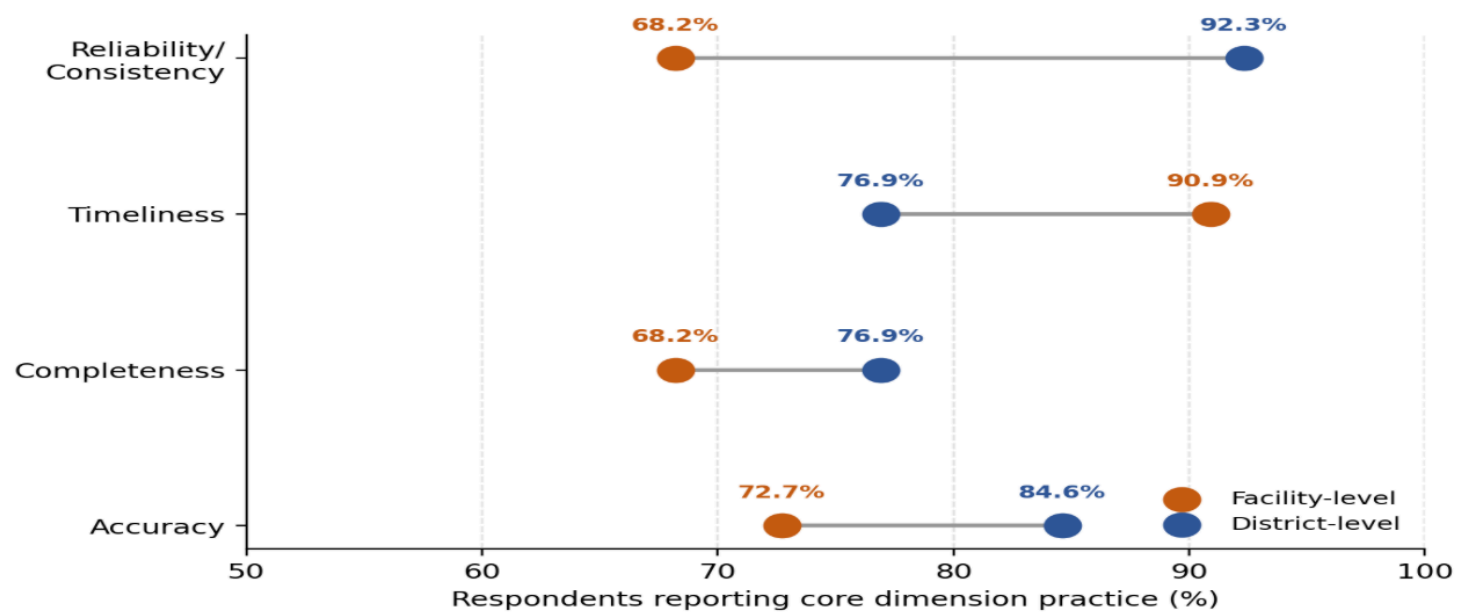


Figure 4

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile1KIIGuideManuscriptonToolsMethodsandRHISDataQuality.docx](#)