

Data Quality and Associated Technological, Organisational, and Behavioural Factors in the Routine Health Information System of Osun State, Nigeria: A Cross-Sectional Study

Sunday Akingbola Omokanye

omokanyeakin@gmail.com

Obafemi Awolowo University

Daisy Ada Oguizu

Rivers State University

Clement Kevin Edet

Rivers State University

Samuel Olutunde Michael

Rivers State University

Research Article

Keywords: Routine Health Information System, data quality, health information technology, organisational behaviour, mediation analysis, Nigeria, sub-Saharan Africa, PRISM framework

Posted Date: October 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10827852/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

Routine Health Information Systems (RHIS) underpin evidence-based health system decision-making, yet data-quality shortfalls remain widely documented across sub-Saharan Africa, including Nigeria. Prior research has established that technical, organisational, and behavioural factors jointly shape RHIS performance, but few studies have tested these factors together, at a common facility, across levels of care, or examined whether organisational influence operates directly or indirectly.

Methods

A cross-sectional, mixed-methods design was used across 22 randomly and purposively sampled primary, secondary, and tertiary healthcare facilities in Osun State, Nigeria. A structured questionnaire, adapted from the Performance of Routine Information System Management (PRISM) framework, was administered to 297 health workers (100% response rate), alongside an observational facility checklist (23 completed checklists across the 22 facilities). Data quality was assessed across four dimensions (accuracy, completeness, timeliness, consistency). Multiple linear regression, with variance inflation factor and residual diagnostics, tested the direct effects of technological, organisational, and behavioural factors on data quality; a supplementary bootstrapped mediation analysis (2,000 resamples) tested whether organisational factors' effect on data quality was direct or indirect.

Results

Self-reported data quality was strongest for timeliness (91.1% of maximum obtainable score) and weakest for accuracy (71.2%), a pattern corroborated by the facility checklist, which found only 21.7% of facilities had fully accurate monthly reports despite 78.3% completeness. Technological factors ($\beta = 0.101$, $p = .008$) and behavioural factors ($\beta = 0.215$, $p = .003$) significantly predicted data quality (adjusted $R^2 = 0.175$); organisational factors showed no significant direct effect ($p = .595$). Mediation analysis showed organisational factors had a significant total effect on data quality ($c = 0.269$, $p < .001$) operating entirely through behavioural factors (indirect effect = 0.167, 95% CI [0.079, 0.271]) and technological factors (indirect effect = 0.058, 95% CI [0.021, 0.109]), with a non-significant residual direct effect ($c' = 0.044$, $p = .471$), indicating full mediation.

Conclusions

RHIS data quality in Osun State is constrained primarily by accuracy rather than completeness or timeliness, and is most directly improved through behavioural and technological investment; organisational reforms remain relevant but act on data quality only insofar as they change staff

behaviour and technology use. Interventions and policies should be designed, and evaluated, against this mediated pathway rather than against organisational compliance alone.

Trial registration:

Not applicable.

Background

Reliable health information underpins effective decision-making, planning, and resource allocation in any health system, and the World Health Organization (WHO) identifies health information systems as one of six essential building blocks of health system performance, alongside service delivery, health workforce, access to essential medicines, financing, and governance [1]. Data quality, understood as the extent to which data are accurate, complete, timely, and consistent, determines whether this building block can perform its function; deficiencies in data quality have been associated with an estimated 71,000 preventable deaths annually in the United States alone from medical errors linked to poor documentation and information flow [2].

Routine Health Information Systems (RHIS), the systems through which health facilities continuously collect, aggregate, and report service-delivery and disease-surveillance data, are central to this concern. A 2024 assessment by the WHO Regional Office for Africa found overall health information system maturity across the African Region to stand at only 58%, despite the proportion of countries using the DHIS2 platform rising from 68% to 91% between 2019 and 2023 [3]. This finding is instructive: the near-universal adoption of a common digital reporting platform has not, on its own, resolved the region's underlying data-quality problem, consistent with the Performance of Routine Information System Management (PRISM) framework's central premise that technical, organisational, and behavioural determinants of RHIS performance must be addressed together [4].

Nigeria illustrates the same pattern nationally. The country's National Health Management Information System (NHMIS), established in 1993 and migrated to the DHIS2 platform in 2010 [5], operates against persistent evidence of facility-level under-reporting, documented at between 10% and 60% depending on the indicator and setting [6], and a WHO-based assessment found facility-reported NHMIS data could be incomplete for as much as 40% of the time [7]. Osun State, the setting for the present study, reflects these national challenges alongside documented local infrastructural and staffing gaps: a 2024 audit of the state's primary healthcare facilities found a 41.19% disparity between assigned and active health workers, and structural deficiencies in half of surveyed facilities [8].

Prior RHIS research, both within Nigeria and across sub-Saharan Africa more broadly, has generally examined data quality and its determinants in aggregate, without distinguishing systematically between facility levels, and has rarely tested technological, organisational, and behavioural determinants together within a single instrument applied across levels of care. This leaves an important gap for policymakers

seeking to target scarce improvement resources: without evidence on which determinants matter most, and at which facility level, interventions risk being poorly matched to the facilities they are meant to help.

This study addresses that gap by assessing the current state of RHIS data quality, in terms of accuracy, completeness, timeliness, and reliability, across primary, secondary, and tertiary healthcare facilities in Osun State, Nigeria, and by examining the technological, organisational, and behavioural factors associated with it. Specifically, the study aimed to: (i) examine the tools and methods of Osun State's RHIS; (ii) assess the current state of RHIS data quality; (iii) explore information technology factors affecting RHIS data quality; (iv) investigate organisational factors affecting RHIS data quality; and (v) explore health workers' behavioural factors affecting RHIS data quality, including testing whether the influence of these three categories of factor on data quality is direct, indirect, or both.

Theoretical framework

Two complementary theories informed this study's design and its interpretation of findings. First, the PRISM framework [4, 9, 10] provided the study's overarching structure. PRISM conceptualises RHIS performance as a chain running from input to impact: technical, organisational, and behavioural determinants (input) shape data collection, transmission, and analysis practices (process), which in turn determine data quality and use (output), health system performance (outcome), and ultimately health outcomes (impact). A defining feature of PRISM, and the reason it was selected over single-factor models, is its explicit treatment of technical, organisational, and behavioural determinants as interacting rather than independent; the framework does not assume that each determinant acts on RHIS output in isolation, leaving open the possibility that one determinant's influence is transmitted through another rather than exerted directly. This logic directly motivated the study's decision to test not only the direct effects of technology, organisational, and behavioural factors on data quality, but also whether organisational factors' effect might operate indirectly through the other two.

Second, the Technology Acceptance Model (TAM) [11] informed the study's expectations regarding the technology–data quality relationship specifically. TAM holds that an individual's use of a technology is shaped by its perceived usefulness and perceived ease of use, both of which are themselves shaped by external factors such as infrastructure, training, and system design. Applied to RHIS, TAM implies that technology's effect on data quality should not be expected to be uniform across settings: where perceived usefulness and ease of use are high, typically in less complex reporting environments, technology should show a stronger direct effect on data quality than in more complex environments where the same technology may be perceived as burdensome or poorly integrated with existing workflows. This distinction motivated the study's decision to test the technology–data quality relationship separately across facility levels rather than only in aggregate.

Conceptual framework

Figure 1 presents the conceptual framework derived from these two theories. Technology, organisational, and behavioural factors are modelled as the study's independent variables, consistent with PRISM's input category, while RHIS data quality, comprising accuracy, completeness, timeliness, and consistency, is modelled as the dependent variable. Technology and behavioural factors are hypothesised to exert a direct effect on data quality (solid arrows), consistent with their status as proximate, facility-level determinants under both PRISM and TAM. Organisational factors are hypothesised to exert both a direct effect (dashed arrow, tested as Hypothesis 2) and an indirect effect transmitted through technology and behavioural factors (curved grey arrows), consistent with PRISM's characterisation of organisational determinants as operating partly through their influence on the technical and behavioural conditions closer to the point of data generation. This dual specification, rather than assuming organisational factors act only directly, is what allowed the study to test, rather than simply assume, how organisational factors relate to data quality in this setting.

Methods

Study design and setting

A cross-sectional, mixed-methods design was used to assess the current state of RHIS data quality and its associated determinants at a single point in time. The study was conducted in Osun State, south-western Nigeria, a state of approximately 4.3 million people [12] organised into three senatorial districts and 30 Local Government Areas, with a health system comprising approximately 1,019 public and private healthcare facilities [13].

Population and sampling

The study population comprised health workers directly involved in the collection, collation, analysis, and reporting of routine health information within Osun State's public healthcare system, including Health Information Management officers, nurses, medical officers, community health officers and extension workers, medical laboratory scientists, and pharmacists. Staff with less than six months of service at their facility were excluded.

A multistage sampling design was used. In the first stage, 22 of the state's 589 public healthcare facilities were selected using stratified and cluster sampling across the three senatorial districts and three facility levels (primary, secondary, tertiary): 14 primary facilities were selected proportionate to each district's share of the state's 532 primary health centres; 6 secondary facilities were apportioned equally (two per district) from the state's 55 secondary facilities; and both of the state's tertiary facilities (Obafemi Awolowo University Teaching Hospitals Complex and Osun State University Teaching Hospital) were included through total enumeration, given their role as apex referral and aggregate data centres. Table 1 presents the resulting sample distribution.

Table 1
Distribution of sampled healthcare facilities by senatorial district
and facility level (N = 22)

Senatorial District	Primary	Secondary	Tertiary	Total
Osun East	6	2	1	9
Osun Central	5	2	1	8
Osun West	3	2	0	5
Total	14	6	2	22

In the second stage, individual respondents were selected from each facility using simple random sampling from a complete sampling frame of eligible staff, using computer-generated random numbers to ensure equal selection probability. Sample size was determined using Cochran's [14] formula for an infinite population ($Z = 1.96$, $p = 0.5$, $e = 0.05$), giving an initial estimate of $n_0 = 385$, corrected for the finite study population ($N = 1,139$ eligible health workers) using Cochran's finite population correction, $n = n_0 / [1 + (n_0 - 1)/N] = 288$, and inflated by a 3% non-response allowance to a final target sample of 297. All 297 distributed questionnaires were completed and retrieved (100% response rate), reflecting direct, researcher-administered delivery with immediate on-site retrieval by five trained research assistants over a three-month period.

Alongside the questionnaire, a structured observational checklist was completed once per facility across 21 of the 22 sampled facilities; at Osun State University Teaching Hospital, two checklists were completed given the scale of the institution, yielding 23 completed checklists in total.

Instruments

The structured questionnaire and the observational checklist were adapted and developed for use in the present study, using the Performance of Routine Information System Management (PRISM) framework [4, 10] which conceptualises routine health information system (RHIS) performance as a function of technical, organisational and behavioural determinants. The structured questionnaire included six sections: socio-demographic characteristics (10 items), RHIS tools and methods (14 items), current data quality (10 items, including 4 dimensions: accuracy, completeness, timeliness and consistency), information technology factors (12 items) and organisational factors (14 items), and health workers' behavioural factors (18 items). Technological, organisational and behavioural determinants were measured on a five-point Likert scale. The observational checklist was also based upon the PRISM framework and included four components that evaluated RHIS tools and methods, data quality dimensions, information technology infrastructure and organisational support systems. It was finalized by the research team after observing the various facilities and checking of relevant RHIS records and documents. The entire structured questionnaire and observation checklist are provided in the Supplementary File 1 and 2 respectively.

Reliability and validity

Instrument reliability and validity were assessed in two stages. A pilot study ($n = 20$) was conducted in neighbouring Oyo State ahead of the main fieldwork; pilot-stage Cronbach's alpha for the Facility Questionnaire was 0.790. Following completion of the main fieldwork, reliability was reassessed on the full analytic sample ($n = 297$): Cronbach's alpha was 0.817 for the composite Data Quality scale, 0.909 for Technology factors, 0.822 for Organisational factors, and 0.763 for Behavioural factors, all exceeding the conventional 0.70 threshold for acceptable reliability. Construct validity was similarly reassessed on the full sample using the Organisational Factors scale: the Kaiser-Meyer-Olkin measure of sampling adequacy was 0.875 (classified as 'meritorious' per Kaiser [15]), Bartlett's Test of Sphericity was significant ($\chi^2 = 1474.14$, $df = 91$, $p < .001$), and principal component analysis with varimax rotation extracted three components with eigenvalues greater than 1, explaining 57.3% of total variance.

Data analysis

Data were analysed using IBM SPSS Statistics version 30.0, supplemented by Python (statsmodels, pingouin) for regression diagnostics and mediation analysis. Descriptive statistics (frequencies, percentages, means, and standard deviations) summarised respondent characteristics and RHIS data quality dimensions. Multiple linear regression tested the effect of technology, organisational, and behavioural composite scores on data quality, both overall and stratified by facility level; variance inflation factors (VIF) assessed multicollinearity, and Shapiro-Wilk and Breusch-Pagan tests assessed residual normality and homoscedasticity, with heteroscedasticity-robust (HC3) standard errors reported where these diagnostics indicated departures from standard ordinary least squares assumptions. A supplementary bootstrapped mediation analysis (2,000 resamples), using organisational factors as the predictor, behavioural and technology factors as parallel mediators, and data quality as the outcome, tested whether organisational factors' effect on data quality operates directly or indirectly. Statistical significance was set at $\alpha = 0.05$ throughout.

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 data collection, participation was voluntary, and all data were coded and stored in secure, encrypted files with restricted access to protect participant confidentiality.

Results

Respondent characteristics

Of the 297 respondents, 74.7% (n = 222) were female. The sample was concentrated in the 30–39-year age group (33.7%), and 47.5% held a bachelor's degree. Respondents were drawn predominantly from primary healthcare facilities (42.7%), followed by secondary (36.4%) and tertiary (20.9%) facilities, consistent with the proportional sampling design. Table 2 presents the full demographic profile, and Fig. 2 shows the facility-category distribution graphically.

Table 2
Socio-demographic characteristics of respondents (N = 297)

Variable	N	%
Sex: Female	222	74.7%
Sex: Male	75	25.3%
Age 30–39 years	100	33.7%
Age 40–49 years	83	27.9%
Age 18–29 years	82	27.6%
Age 50 + years	32	10.7%
Bachelor's degree	141	47.5%
HND	63	21.2%
OND/ND	63	21.2%
Postgraduate degree	30	10.1%
Married	202	68%
Single	91	30.6%
Health Information Management officer	84	28.3%
Nurse	84	28.3%
Community Health Officer	36	12.1%
Medical Laboratory Scientist	32	10.8%
Community Health Extension Worker	22	7.4%
Pharmacist	14	4.7%
Medical Officer	10	3.4%
Primary facility	127	42.7%
Secondary facility	108	36.4%
Tertiary facility	62	20.9%

RHIS tools and methods

Standardised Ministry of Health registers were widely used and self-reported: 60.6% of respondents strongly agreed their facility used standard registers, and observational checklist data confirmed this, with 87.0% of facilities found to have functioning MOH registers. Reporting relied heavily on a hybrid paper-electronic model: 57.6% of respondents strongly agreed that paper-based registers remained the primary source of data prior to digital entry, and the observational checklist found 82.6% of facilities entering paper-recorded data into DHIS2 or submitting it to the Local Government Area. Integrated Disease Surveillance and Response (IDSR) tools were in use at 60.9% of facilities per the observational checklist, though only 39.1% had documented use of locally modified forms when standard tools were unavailable, indicating both reasonable tool coverage and continued reliance on manual processes.

Current state of RHIS data quality

Self-reported data quality varied markedly across its four dimensions. Expressed as a percentage of each dimension's maximum obtainable score, timeliness was highest (91.1%; mean = 9.11, SD = 1.7, of a 10-point maximum), followed by completeness (87.8%; mean = 8.78, SD = 2.0, of a 10-point maximum) and consistency (86.9%; mean = 13.04, SD = 3.2, of a 15-point maximum), while accuracy was comparatively weakest (71.2%; mean = 10.68, SD = 4.3, of a 15-point maximum). Figure 3 presents this comparison graphically.

This pattern was corroborated, and considerably sharpened, by the observational facility checklist, which assessed accuracy more objectively through direct register-to-summary comparison rather than staff self-report. Only 21.7% of the 23 checklists found facility monthly reports to be fully accurate, with substantial register-summary mismatches for malaria (52.2% mismatched) and antenatal care (39.1% mismatched) data, and missing data exceeding 65% for immunisation, measles, nutrition, diarrhoea, tuberculosis, hypertension, upper respiratory tract infection, and diabetes mellitus indicators. By contrast, completeness was confirmed as comparatively strong (78.3% of facilities with complete data), and consistency and reliability practices were similarly strong, with 95.7% of facilities monitoring data over time and 95.7% possessing a documented data-collection procedure manual. Timeliness showed a more divided picture under objective assessment: while 73.9% of facilities submitted reports within the first week of the following month, only approximately 45% of reports were filed strictly on time in the two audited months (January and July 2024), with a substantial proportion of reporting data missing for those periods. Table 3 summarises these findings.

Table 3

Current state of RHIS data quality: self-reported (n = 297) and observational checklist (n = 23) findings

Dimension	Self-report mean score (% of maximum)	Checklist finding	Interpretation
Accuracy	10.68/15 (71.2%)	21.7% fully accurate; >65% missing data for 8 of 12 tracer indicators	Weakest dimension; objective assessment reveals a substantially larger gap than self-report
Completeness	8.78/10 (87.8%)	78.3% of facilities with complete data	Consistently strong across both measures
Timeliness	9.11/10 (91.1%)	73.9% within first week; ~45% strictly on-time in audited months	Strong self-report; objective audit shows more moderate performance
Consistency/Reliability	13.04/15 (86.9%)	95.7% monitor data over time; 95.7% have procedure manuals	Consistently strong across both measures

Technological, organisational, and behavioural factors

Technology factors showed the lowest overall attainment of the three determinant categories (mean = 35.34, SD = 10.99, of a 60-point maximum; 58.9% of maximum), with 53.2% of respondents reporting computer use for report generation and only 43.5% reporting reliable internet access. The observational checklist found similarly limited infrastructure: only 26.1% of facilities had functioning computer systems and 21.7% had internet access, though 82.6% used a Health Management Information System of some form. Organisational factors were more favourably rated (mean = 54.57, SD = 8.15, of a 70-point maximum; 78.0% of maximum), with 86.2% of respondents agreeing or strongly agreeing that supervisors provided supportive supervision. Behavioural factors were rated most favourably of all three (mean = 72.09, SD = 9.85, of a 90-point maximum; 80.1% of maximum), with 91.2% of respondents agreeing or strongly agreeing that they were punctual and 87.2% agreeing that collecting information was meaningful to them.

Data quality across facility levels

Overall data quality was highest in primary healthcare facilities (mean = 43.59, SD = 5.78), followed by tertiary (mean = 40.31, SD = 7.73) and secondary facilities (mean = 40.03, SD = 7.48). This pattern was not uniform across determinants, however: tertiary facilities recorded the highest technology (mean = 41.26, SD = 8.60) and organisational (mean = 56.15, SD = 5.82) scores, while primary facilities recorded the highest behavioural score (mean = 73.34, SD = 10.27). Table 4 presents the full facility-level breakdown across all data-quality sub-dimensions and determinant sub-components, and Fig. 4 presents the four composite scores by facility level as a percentage of each determinant's own maximum obtainable score.

Table 4
Comparison of data quality and associated factors across facility levels (mean $\pm$ SD)

Assessed parameter	Primary (mean $\pm$ SD)	Secondary (mean $\pm$ SD)	Tertiary (mean $\pm$ SD)
Data accuracy	11.31 $\pm$ 3.72	9.78 $\pm$ 3.78	10.95 $\pm$ 3.64
Data completeness	9.13 $\pm$ 1.44	8.68 $\pm$ 1.85	8.27 $\pm$ 2.05
Data timeliness	9.25 $\pm$ 1.53	9.10 $\pm$ 1.48	8.85 $\pm$ 1.74
Data consistency	13.91 $\pm$ 1.67	12.49 $\pm$ 3.31	12.23 $\pm$ 2.98
Computer facilities	9.21 $\pm$ 3.56	9.42 $\pm$ 3.73	11.76 $\pm$ 2.67
Internet facilities	12.63 $\pm$ 4.70	11.27 $\pm$ 4.50	14.03 $\pm$ 3.60
HMIS facilities	9.28 $\pm$ 3.89	7.98 $\pm$ 3.76	11.89 $\pm$ 3.32
Data entry facilities	6.88 $\pm$ 2.37	6.86 $\pm$ 2.56	7.77 $\pm$ 1.89
Decision-making style	13.67 $\pm$ 3.27	13.65 $\pm$ 3.36	14.03 $\pm$ 1.97
Management style	41.40 $\pm$ 6.56	39.43 $\pm$ 6.43	42.11 $\pm$ 5.19
Motivation	21.57 $\pm$ 3.40	21.17 $\pm$ 3.15	22.03 $\pm$ 3.09
Competence	33.71 $\pm$ 6.78	31.63 $\pm$ 5.67	32.81 $\pm$ 4.87
Attitude	14.13 $\pm$ 2.75	14.00 $\pm$ 3.30	13.58 $\pm$ 2.66
Data Quality (composite)	43.59 $\pm$ 5.78	40.03 $\pm$ 7.48	40.31 $\pm$ 7.73
Technology factors (composite)	34.94 $\pm$ 10.93	32.41 $\pm$ 11.05	41.26 $\pm$ 8.60
Organisational factors (composite)	55.08 $\pm$ 8.61	53.07 $\pm$ 8.57	56.15 $\pm$ 5.82
Behavioural factors (composite)	73.34 $\pm$ 10.27	70.55 $\pm$ 10.08	72.19 $\pm$ 8.24

The facility-level breakdown clarifies the composition of this complementary pattern. At the sub-dimension level, primary facilities led on data accuracy, completeness, timeliness, and consistency alike, and also led on decision-making style, motivation, and attitude; their comparative weakness was confined to the technology sub-components (computer, internet, HMIS, and data-entry facilities), all of which were highest at tertiary level. Secondary facilities scored lowest of the three levels on almost every sub-dimension, consistent with their intermediate-to-lowest ranking on the overall composite scores. This finer-grained picture reinforces the interpretation offered above: tertiary facilities' technological and, to a lesser extent, organisational-management advantage does not extend to the behavioural and data-quality sub-dimensions that this study's regression and mediation results identify as the more proximate determinants of data quality.

Predictors of data quality

Multiple linear regression tested the effect of technology, organisational, and behavioural factors on data quality. Technology factors ($\beta = 0.101$, $p = .008$) and behavioural factors ($\beta = 0.215$, $p = .003$) were significant positive predictors; organisational factors were not significant ($\beta = 0.044$, $p = .595$). The model explained a modest 17.5% of variance in data quality (adjusted $R^2 = 0.175$). Variance inflation factors for all three predictors (1.19–1.76) indicated no meaningful multicollinearity. Residual diagnostics indicated departures from normality (Shapiro-Wilk $W = 0.937$, $p < .001$) and homoscedasticity (Breusch-Pagan LM = 24.53, $p < .001$), both common in bounded Likert-sum outcomes of this kind; heteroscedasticity-robust (HC3) standard errors, reported in Table 5, did not alter the pattern of significance. Facility-level stratification showed a consistent role for behavioural factors (significant at secondary, $p = .001$, and marginal at tertiary level, $p = .052$) and a role for technology factors concentrated at the primary level ($p = .009$); organisational factors were not a significant direct predictor at any facility level.

Table 5
Multiple regression of technology, organisational, and behavioural factors on RHIS data quality ($n = 297$), with robust standard errors

Predictor	B	Robust SE	z	p
(Constant)	19.810	4.876	4.06	.000
Technology factors	0.101	0.038	2.65	.008
Organisational factors	0.044	0.083	0.53	.595
Behavioural factors	0.215	0.073	2.94	.003

Adjusted $R^2 = 0.175$. Variance inflation factors: Technology = 1.19, Organisational = 1.76, Behavioural = 1.70 (no multicollinearity). Dependent variable: composite Data Quality score.

Mediation analysis: resolving the organisational factors result

Because organisational factors correlated more strongly with behavioural factors ($r = 0.63$) than with data quality directly ($r = 0.31$), a bootstrapped mediation analysis (2,000 resamples) tested whether organisational factors' influence on data quality operates indirectly, through behavioural and technology factors, rather than directly. Organisational factors had a significant total effect on data quality ($c = 0.269$, $p < .001$). Once behavioural and technology factors were entered as parallel mediators, the direct effect of organisational factors fell to non-significance ($c' = 0.044$, $p = .471$), while both indirect pathways remained significant: via behavioural factors (indirect effect = 0.167, 95% CI [0.079, 0.271]) and via technology factors (indirect effect = 0.058, 95% CI [0.021, 0.109]). Because both indirect-effect confidence intervals excluded zero while the direct effect did not differ significantly from zero, this

constitutes full (complete) mediation. Figure 5 illustrates the mediation pathway, and Table 6 reports the full path coefficients.

Table 6

Mediation analysis: organisational factors to (behavioural + technology factors) to data quality

Path	Coefficient	p / 95% CI
Organisational → Behavioural factors (a1)	0.776	p < .001
Organisational → Technology factors (a2)	0.575	p < .001
Behavioural factors → Data Quality, controlling for Technology (b1)	0.235	p < .001
Technology factors → Data Quality, controlling for Behavioural (b2)	0.106	p = .001
Total effect: Organisational → Data Quality (c)	0.269	p < .001
Direct effect: Organisational → Data Quality, both mediators controlled (c')	0.044	p = .471
Indirect effect via Behavioural factors	0.167	95% CI [0.079, 0.271]
Indirect effect via Technology factors	0.058	95% CI [0.021, 0.109]

In summary, organisational factors are significantly related to RHIS data quality in Osun State, but this relationship is fully explained by their influence on staff behaviour and technology use, rather than by any independent direct effect.

Discussion

This study set out to assess RHIS data quality in Osun State, Nigeria, and to determine the technological, organisational, and behavioural factors associated with it across facility levels. Three findings merit particular discussion: the divergence between accuracy and the other three data-quality dimensions; the facility-level complementarity between infrastructure and behaviour; and the resolution, through mediation analysis, of organisational factors' apparently contradictory relationship with data quality.

Accuracy as the binding constraint on RHIS data quality

Completeness, timeliness, and consistency were consistently rated strong, both by self-report and by observational checklist, while accuracy lagged substantially behind on both measures, and the gap widened considerably once accuracy was assessed objectively through register-to-summary comparison rather than staff perception. This pattern is consistent with evidence from Ghana and Ethiopia showing that completeness is often the most easily enforced dimension of data quality because it is directly visible to supervisors, while accuracy requires active verification that is more easily neglected under

workload pressure [16, 17, 18]. The implication is that RHIS strengthening efforts in Osun State, and plausibly in comparable Nigerian and sub-Saharan African settings, should not assume that high completeness or timeliness indicates high overall data quality; targeted register-to-report reconciliation and routine data-quality audits are needed specifically to close the accuracy gap that this study's dual-method design was able to detect but that a self-report-only design would likely have understated.

Complementary, not uniform, facility-level strengths

Primary facilities recorded the highest overall data quality despite the lowest technology scores, while tertiary facilities recorded the strongest technological and organisational infrastructure but only intermediate data quality. This complementary pattern suggests that behavioural and organisational cohesion in smaller, more directly supervised primary facilities can substitute, to a degree, for limited technological capacity, consistent with findings from Ethiopia and South Africa that well-supervised, less complex facility teams often outperform larger, technically better-resourced institutions on routine reporting discipline [19, 20]. Tertiary facilities' comparatively weaker data quality despite stronger infrastructure similarly echoes evidence that technological capacity alone does not resolve data-quality problems introduced by higher patient volumes, more complex workflows, and more decentralised documentation [21]. This implies that facility-level RHIS interventions in Osun State should be differentiated rather than uniform: infrastructure investment is likely to yield the greatest marginal benefit at the primary level, where behavioural strengths already exist to build on, while tertiary facilities require process-integration and workflow-simplification interventions more than further hardware investment.

The facility-level breakdown in Table 4 shows that primary facilities' technology disadvantage is not spread evenly across its sub-components. The gap between primary and tertiary facilities was largest for computer facilities (9.21 vs. 11.76, a 27.7% shortfall) and HMIS facilities (9.28 vs. 11.89, a 28.1% shortfall), but comparatively narrow for internet connectivity (12.63 vs. 14.03, an 11.1% shortfall) and, most notably, for data-entry facilities (6.88 vs. 7.77, a 12.9% shortfall). In other words, primary facilities' technology deficit is concentrated in hardware and platform availability – functioning computers and an accessible HMIS – rather than in the data-entry practices staff apply once a system is in front of them. This distinction sharpens the intervention implication drawn above: because the practice-level component of technology use is already comparatively strong at primary facilities, closing the primary–tertiary technology gap is plausibly more a matter of provisioning hardware and platform access than of retraining staff on data-entry technique, whereas tertiary facilities' comparable data-entry scores despite markedly better infrastructure reinforce the earlier point that infrastructure alone does not translate into stronger reported data-entry practice once workflow complexity is taken into account.

Resolving the organisational factors result through mediation

The most analytically significant finding of this study is the resolution of an apparent contradiction between the direct-effects regression, in which organisational factors showed no significant association

with data quality, and the wider RHIS literature, which consistently identifies organisational support, supervision, and governance as important determinants of RHIS performance [22, 23, 24]. Rather than treating the non-significant direct coefficient as evidence that organisational factors do not matter in this context, this study tested the pathway suggested by the correlation structure of the data directly, and found full mediation: organisational factors have a significant total effect on data quality, operating entirely through their influence on staff behaviour and, to a lesser extent, technology use. This finding is consistent with, and provides a specific statistical mechanism for, prior arguments that organisational determinants of RHIS performance often act indirectly [25, 26], and offers other RHIS researchers encountering a similarly non-significant direct organisational effect a concrete analytical strategy, testing indirect pathways through behavioural and technological mediators, before concluding that organisational factors are unimportant.

Practically, this finding has a direct implication for intervention design: organisational reforms aimed at improving RHIS data quality, whether new supervisory structures, governance changes, or updated reporting frameworks, are likely to succeed only to the extent that they change how motivated, competent, and technologically equipped frontline staff become. An organisational intervention that strengthens formal structures without altering staff behaviour or technology access should not be expected, on the basis of this study's findings, to improve data quality on its own.

Comparison with the wider literature

The significant direct effects of technology and behavioural factors on data quality are consistent with evidence from Rwanda, where electronic health record functionality significantly predicted data completeness and timeliness [21], and from South Sudan and Ethiopia, where staff motivation and competence were identified as leading predictors of RHIS performance even under resource constraints [26, 27]. Similarly, in Somalia and Ghana, unstable infrastructure and low technology utilisation were found to depress RHIS reporting quality even where formal systems existed on paper [28, 29], reinforcing this study's finding that technology's effect on data quality is conditional on actual, reliable use rather than nominal availability. The behavioural mechanism identified here also echoes evidence from Senegal, where health workers' perceptions of RHIS tasks as meaningful, rather than burdensome, were found to be a stronger predictor of data-use behaviour than formal training exposure alone [30], a pattern this study's high self-reported behavioural scores (80.1% of maximum) are broadly consistent with.

The modest overall variance explained (17.5–18.9%, depending on model specification) is itself an informative finding: it indicates that unmeasured factors, plausibly including funding, workload, patient volume, and factors outside the PRISM framework's technical-organisational-behavioural triad [9], also shape RHIS data quality in this setting, and that technological and behavioural investment, while empirically justified, should not be expected to resolve data-quality gaps on its own. Evidence from Ethiopia similarly found that data-quality audit and feedback interventions produced measurable but partial improvements in accuracy and completeness [31], consistent with this study's implication that no single determinant category is likely to be sufficient on its own to close the accuracy gap identified here.

Interpreting the findings through PRISM and the Technology Acceptance Model

The findings reported above map closely onto the two theories that informed the study's conceptual framework (Fig. 1), and each theory helps explain a specific empirical pattern rather than the findings merely being consistent with theory in general terms.

The full mediation result for organisational factors is, in effect, a direct empirical confirmation of PRISM's input-process-output structure [4]. PRISM does not model organisational determinants as acting on RHIS output independently of technical and behavioural determinants; rather, it models all three as inputs that jointly shape the process of data collection, transmission, and use, through which output, in this case, data quality, is ultimately produced. A non-significant direct effect of organisational factors on data quality, occurring alongside a significant total effect operating through technology and behavioural factors, is exactly the pattern PRISM's input-process logic would predict for an input variable whose influence is transmitted through the process layer rather than exerted at the output layer directly. This study appears to be among the first in the Nigerian RHIS literature to test, rather than assume, this specific structural implication of PRISM using formal mediation analysis, and the result provides empirical support for treating PRISM's tripartite determinants as an interacting system rather than as three independent, additive predictors, a distinction with direct consequences for how organisational interventions should be designed, as discussed above.

The Technology Acceptance Model [11] similarly helps explain a pattern this study would otherwise be difficult to interpret: technology factors' direct effect on data quality was significant at the primary facility level but not at secondary or tertiary level, despite tertiary facilities having markedly higher technology scores overall. TAM attributes technology's effect on outcomes to perceived usefulness and perceived ease of use rather than to infrastructure availability alone; where reporting workflows are simpler and more direct, as in primary facilities, the same or even lesser technological capacity may be perceived as more useful and easier to use, translating more directly into improved data quality, while in the more complex, multi-programme reporting environments typical of tertiary facilities, greater technological capacity does not guarantee correspondingly higher perceived usefulness or ease of use. This is consistent with this study's finding that tertiary facilities' technology and organisational advantages did not translate into a commensurate data-quality advantage over primary facilities, and it implies that technology investment decisions in Osun State's RHIS should be paired with facility-specific workflow simplification, particularly at higher-complexity facility levels, rather than treated as a uniform, infrastructure-only intervention.

Implications for policy, practice, and research

The findings carry implications for policy, practice, and future research. For policy, the mediation finding implies that RHIS strengthening strategies and budgets framed primarily around organisational or governance reform should be redesigned to specify behavioural and technological outcomes explicitly,

since this study found organisational investment reaches data quality only through these channels. For practice, the accuracy-specific gap identified through dual self-report and observational assessment implies that facility-level quality assurance activities should move beyond monitoring reporting completeness and timeliness, both already strong, toward active, routine verification of reported values against source records. For research, the study demonstrates the analytical value of testing indirect pathways when a theoretically expected direct effect is not found in a standard regression model; researchers examining organisational determinants of health information system performance elsewhere may find a similar mediated, rather than absent, relationship.

Strengths and limitations

This study's principal strengths are its dual self-report and independent observational assessment of data quality, which revealed an accuracy gap a single-method design would likely have understated; its explicit stratification by facility level; and its use of formal mediation analysis to test, rather than assume, how organisational factors relate to data quality. Several limitations should be noted. Its cross-sectional design permits identification of statistical association, including the mediation pathway, but not causal inference; a longitudinal or experimental design would be required to confirm that organisational change causally precedes the behavioural and technological change with which it is associated here. Self-reported data on motivation, supervision, and data-handling practices may be subject to social desirability bias, although this risk was partly mitigated by the inclusion of an independent observational checklist. The observational checklist itself covered a defined set of tracer health programmes and a limited time window, and may not generalise to all RHIS service areas or reporting periods. The study was conducted within 22 purposively and randomly sampled facilities in a single Nigerian state, which may limit generalisability to other states or health system contexts, and five of the nine health-data-reporting zones examined in supplementary sub-analyses had fewer than 15 respondents, too few to support reliable regression, limiting the granularity of zone-level conclusions.

Conclusions

This study found that RHIS data quality in Osun State, Nigeria is constrained primarily by accuracy rather than by completeness, timeliness, or consistency, a pattern evident in both self-reported data and, more starkly, in objective observational checklist findings that showed only 21.7% of facilities produced fully accurate monthly reports despite strong completeness and reporting-culture indicators. Technological and behavioural factors were significant direct predictors of data quality, with behavioural factors, staff motivation, competence, and attitude, emerging as the single strongest determinant. Organisational factors showed no significant direct effect in the regression models, but a supplementary mediation analysis established that organisational factors do have a significant total relationship with data quality, one that operates entirely through their influence on staff behaviour and technology use rather than through any independent channel. Taken together, these findings indicate that improving RHIS data quality in Osun State, and in comparable low-resource, sub-national health systems, requires prioritising accuracy-specific verification mechanisms alongside targeted behavioural and technological investment,

with organisational reform treated as an upstream enabler of these two channels rather than as a direct, independent lever.

Recommendations

Based on the findings, 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 institutionalise routine, facility-level register-to-summary reconciliation as a standing data-quality assurance activity, given that this study's principal finding was a substantial accuracy gap invisible to self-report but clearly detectable through direct record comparison. Health Information Management officers and Monitoring and Evaluation officers at each facility should conduct and document a monthly register-to-DHIS2 reconciliation check, with Local Government Health Departments responsible for verifying compliance through quarterly data-quality audits.
2. The Osun State Ministry of Health, in partnership with development partners and the state's ICT agency, should prioritise technology investment — functional computers, reliable internet connectivity, and electronic health record access — at primary healthcare facilities first, since this study found technology's direct effect on data quality was concentrated at the primary level and primary facilities currently have the lowest technology scores despite the highest overall data quality. The Ministry's Planning, Research and Statistics Department should be responsible for procurement and rollout, with facility heads responsible for utilisation monitoring.
3. Facility heads and Human Resource Departments within Osun State's healthcare institutions should design and implement structured, non-financial staff recognition and feedback mechanisms specifically targeting RHIS data management performance, given that behavioural factors were confirmed as this study's strongest direct predictor of data quality and the principal channel through which organisational support translates into results. This should include quarterly data-quality performance feedback to individual staff and facility-level recognition of accurate, complete, and timely reporting.
4. The Osun State Ministry of Health's policy and planning leadership should redesign organisational RHIS interventions — supervisory frameworks, governance structures, and reporting policy revisions — to explicitly target staff behaviour change and technology adoption as intermediate outcomes, rather than treating organisational compliance as a sufficient objective in itself, since this study found organisational factors influence data quality entirely through these two channels. Monitoring and Evaluation units should track behavioural and technology-use indicators, not only structural compliance indicators, when evaluating organisational reforms.
5. Health Information Management training institutions in Osun State, and the Federal and State Ministries of Health more broadly, should incorporate accuracy-specific data verification skills and register-reconciliation techniques into pre-service and in-service Health Information Management curricula, given that accuracy, not completeness or timeliness, was this study's most consistently

identified data-quality weakness. Training institutions should be responsible for curriculum design, with the State Ministry of Health responsible for mandating periodic refresher training as a condition of continued RHIS duties.

Contributions to knowledge

This study contributes to the RHIS literature in four specific ways. First, it provides facility-level, rather than aggregate, evidence that primary healthcare facilities can outperform better-resourced secondary and tertiary facilities on RHIS data quality, challenging the assumption that facility level and data quality are positively associated. Second, it identifies behavioural factors, rather than technological or organisational factors, as the strongest direct determinant of RHIS data quality in this setting, while showing this effect explains only a modest share of total variance, indicating other unmeasured determinants remain to be identified. Third, and most significantly, it provides statistical confirmation, through formal mediation analysis, that organisational factors' relationship with RHIS data quality in a sub-national Nigerian setting is real but fully indirect, operating through staff behaviour and technology use, reconciling an apparent contradiction between a non-significant direct-effects regression result and the wider literature's consistent emphasis on organisational determinants, and offering a transferable analytical approach for similar apparent contradictions in future RHIS research. Fourth, it demonstrates the value of pairing self-reported and independently observed data-quality measures within a single RHIS study, since the two methods produced materially different pictures of accuracy specifically, a divergence a single-method study would not have detected.

Abbreviations

CI

Confidence interval

DHIS2

District Health Information Software, version 2

HC3

Heteroscedasticity-consistent standard errors (version 3)

IDSR

Integrated Disease Surveillance and Response

KMO

Kaiser-Meyer-Olkin (measure of sampling adequacy)

MOH

Ministry of Health

NHMIS

National Health Management Information System

PRISM

Performance of Routine Information System Management

RHIS

Routine Health Information System

SD

Standard deviation

TAM

Technology Acceptance Model

VIF

Variance inflation factor

WHO

World Health Organization

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 [author to insert specific approval/reference numbers]. All participants provided written informed consent prior to participation. The study was conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable. This manuscript does not contain any individual person's data in any form (e.g., individual details, images, or videos).

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; it was self-funded by the researchers.

Author Contribution

SAO conceived the study, initiated the study design, participated in data analysis and coordination, and drafted the manuscript. DAO contributed to the study design and coordination and critically reviewed the manuscript for intellectual content. CKE also contributed to the study design and coordination and reviewed the manuscript. SOM participated in the study design and coordination and was actively involved in data collection and data analysis. All authors read and approved the final manuscript.

Acknowledgement

The authors thank the management and staff of the twenty-two participating healthcare facilities in Osun State for their cooperation during data collection, the five trained research assistants who conducted fieldwork, and the Osun State Ministry of Health for facilitating facility access.

Data Availability

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

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. Daneshkohan A. Data quality and data use in primary health care: a case study from Iran. *Inf Med Unlocked*. 2022;28:100855.
3. World Health Organization Regional Office for Africa. Health information system [Internet]. 2024. Available from: <https://www.afro.who.int/health-topics/health-information-system>
4. Hotchkiss DR. Evaluation of the Performance of Routine Information System Management (PRISM) framework: evidence from Uganda. *BMC Health Serv Res*. 2020;10:188.
5. Federal Ministry of Health. NHMIS annual reports [Internet]. 2023. Available from: <https://fmohconnect.gov.ng/nhmis-annual-reports/>
6. 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.
7. Frontiers in Digital Health. Data quality assessment of Nigeria's national health management information system. *Front Digit Health*. 2022.
8. MyPHC/ORO Data Hub. Osun State's primary healthcare system report 2024 [Internet]. 2024. Available from: <https://www.orodata.org/myphc-report-osun-states-primary-healthcare-system-report-2024/>
9. Dehnavieh R, Haghdoost A, Khosravi A, Hoseinabadi F, Rahimi H, Poursheikhali A, et al. The District Health Information System (DHIS2): a literature review and meta-synthesis of its strengths and operational challenges based on the experiences of 11 countries. *Health Inf Manag J*. 2019;48(2):62–75.
10. USAID, Evaluation MEASURE. Using PRISM to strengthen and evaluate health information systems (FS-07-18). Chapel Hill: USAID; 2008.

11. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 1989;13(3):319–40.
12. National Population Commission. Nigeria population census report. Abuja: National Population Commission; 2001.
13. Osun State Government. Strategic health development plan (2010–2015). Osogbo: Osun State Ministry of Health; 2010.
14. Cochran WG. Sampling techniques. 3rd ed. New York: Wiley; 1977.
15. Kaiser HF. An index of factorial simplicity. *Psychometrika.* 1974;39(1):31–6.
16. 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.
17. Syed R, Eden R, Makasi T, Chukwudi I, Mamudu A, Kamalpour M, et al. Digital health data quality issues: systematic review. *J Med Internet Res.* 2023;25:e42615.
18. 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.
19. Moloko SM, Ramukumba MM. Healthcare providers' views of factors influencing family planning data quality in Tshwane District, South Africa. *Afr J Prim Health Care Fam Med.* 2022;14(1):1–10.
20. Cheburet SK, Odhiambo-Otieno GW. Organizational factors affecting data quality of routine health management information system quality: case of Uasin Gishu County Referral Hospital, Kenya. *Int Res J Public Environ Health.* 2016;3(9):201–8.
21. Fraser HSF, Mugisha M, Bacher I, Ngenzi JL, Seebregts C, Umubyeyi A, et al. 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 Surveill.* 2024;10:e49127.
22. Glette MK, Wiig S. The role of organizational factors in how efficiency-thoroughness trade-offs potentially affect clinical quality dimensions: a review of the literature. *Int J Health Gov.* 2021;26(3):250–65.
23. Lemma S, Janson A, Persson LA, Wickremasinghe D, Källestål C. Improving quality and use of routine health information system data in low- and middle-income countries: a scoping review. *PLoS ONE.* 2020;15(10):e0239683.
24. Kawakyu N, Inguane C, Fernandes Q, Gremu A, Floriano F, Manaca N, et al. Determinants of translating routine health information system data into action in Mozambique: a qualitative study. *BMJ Glob Health.* 2024;9:e014970.
25. Ubalaeze E. The impact of digital transformation on healthcare delivery in Nigeria: challenges and recommendations. *Int J Res Sci Innov.* 2024;11(15).
26. Morris LD, Nyongesa MW, Sokiri TD. Factors influencing data quality in routine health information systems in Maridi county, South Sudan. *S Afr J Inf Manag.* 2024;26(1):a1856.

27. Haftu B, Taye G, Ayele W, Habtamu T, Biruk E. A mixed-methods assessment of routine health information system (RHIS) data quality and factors affecting it, Addis Ababa City Administration, Ethiopia, 2020. *Ethiop J Health Dev.* 2021;35(1):15–24.

28. Osman NH. Data quality of routine health information systems and associated factors among public health facilities in Banadir region, Somalia: a cross-sectional study. *Health Inf J.* 2026;32(1):1–16.

29. Mensah NK, Adzakpah G, Kissi J, Abdulai K, Taylor-Abdulai H, Johnson SB, et al. Health professionals' perceptions of electronic health records system: a mixed method study in Ghana. *BMC Med Inf Decis Mak.* 2024;24:254.

30. Muhoza P, Saleem H, Faye A, Tine R, Diaw A, Kante AM, et al. Behavioral determinants of routine health information system data use in Senegal: a qualitative inquiry based on the integrated behavioral model. *Glob Health Sci Pract.* 2022;10(3):e2100686.

31. Endriyas M, Alano A, Mekonnen E, Ayele S, Kelaye T, Shiferaw M, et al. Understanding performance data: health management information system data accuracy in Southern Nations Nationalities and People's Region, Ethiopia. *BMC Health Serv Res.* 2019;19(1):175.

Figures

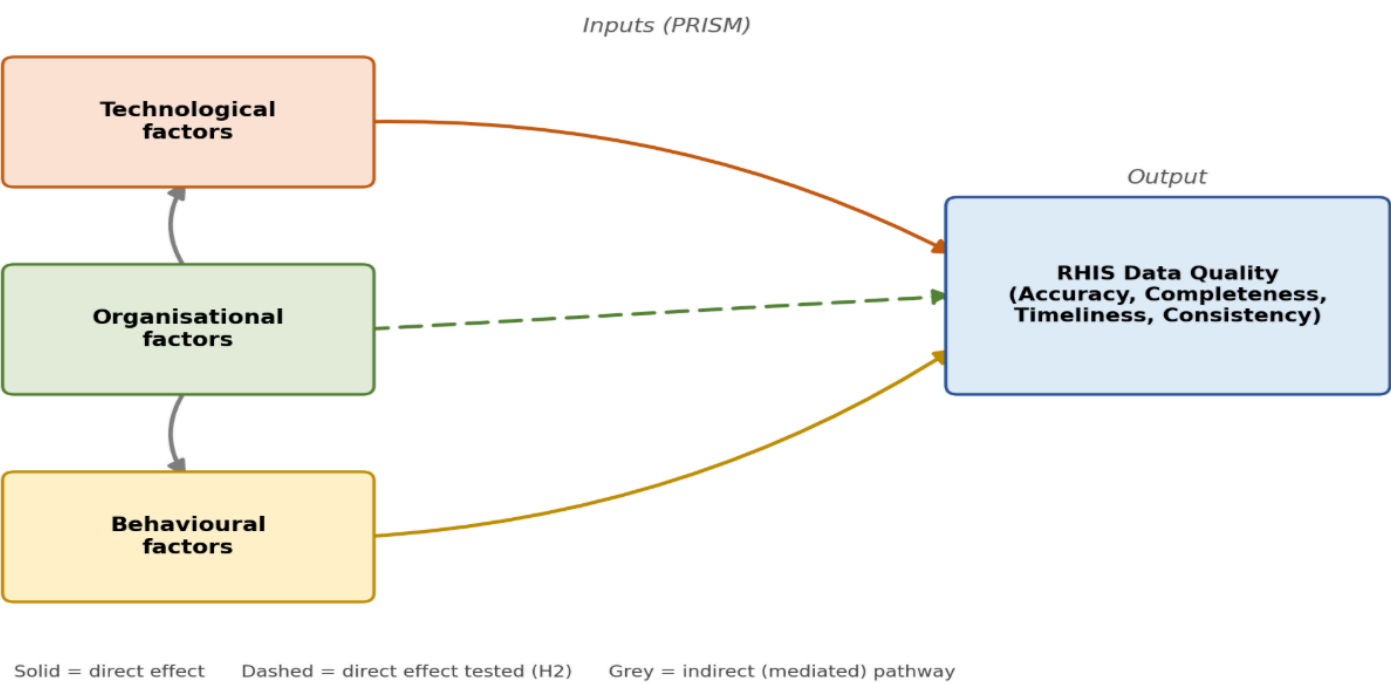


Figure 1

Conceptual framework of the study, derived from the PRISM framework and the Technology Acceptance Model

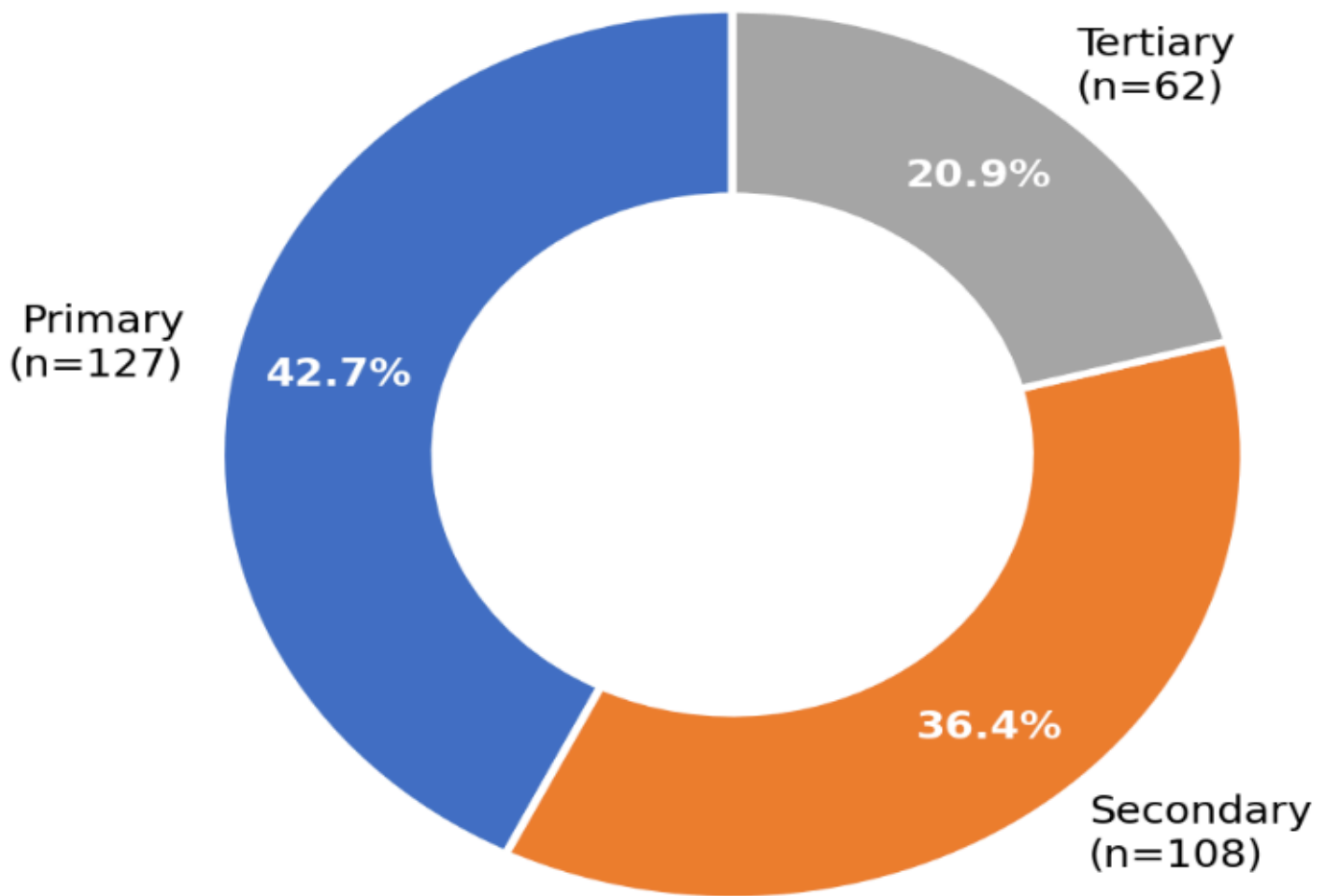


Figure 2

Distribution of respondents by facility category (n = 297)

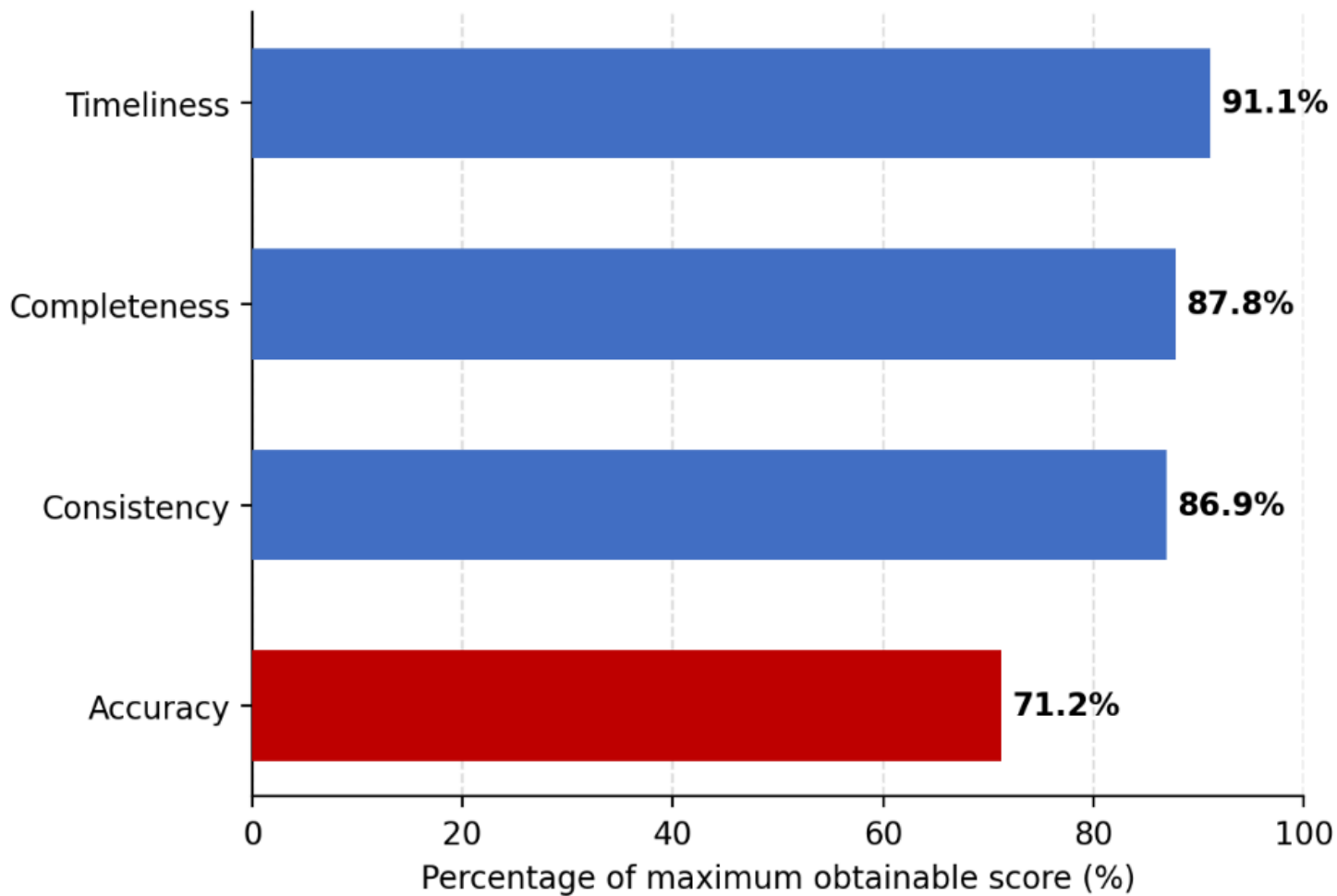


Figure 3

RHIS data quality by dimension, self-reported (n = 297), expressed as a percentage of maximum obtainable score

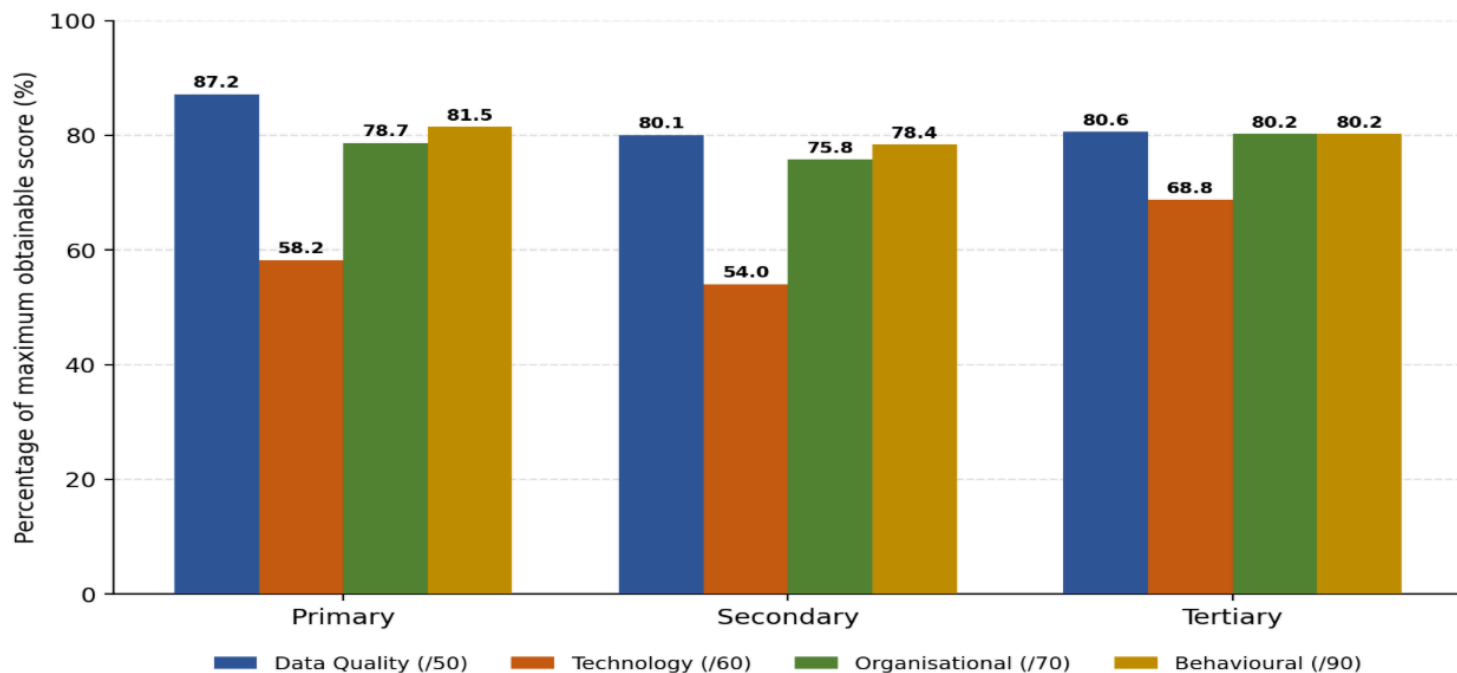
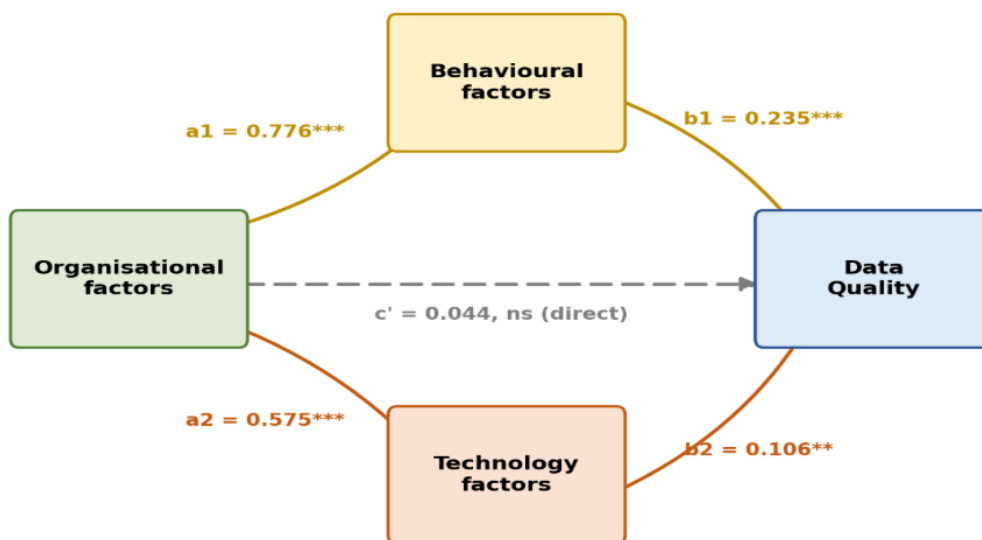


Figure 4

Data quality and associated factors by facility level, expressed as a percentage of each composite measure's maximum obtainable score

Total effect: Organisational → Data Quality, $c = 0.269$, $p < .001$ (full mediation: indirect effects significant, direct effect ns)



*** $p < .001$ ** $p = .001$ Indirect effect via Behavioural = 0.167, 95% CI [0.079, 0.271] Indirect effect via Technology = 0.058, 95% CI [0.021, 0.109]

Figure 5

Mediation of organisational factors' effect on data quality via behavioural and technology factors

Supplementary Files

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

- [SupplementaryIforRHISDataQualityManuscript.docx](#)
- [SupplementaryIIRHISDataQualityManuscript.docx](#)