HAWASSA UNIVERSITY COLLEGE OF MEDICINE AND HEALTH SCIENCES SCHOOL OF PUBLIC HEALTH

THE LEVEL OF ROUTINE HEALTH INFORMATION SYSTEM DATA QUALITY AND ASSOCIATED FACTORS AT PUBLIC HOSPITAL IN SILTE ZONE, CENTRAL REGION, ETHIOPIA, 2023

BY: NIGUSSIE DUKAMO (BSC.NURSING)

NOVEMBER, 2023

HAWASSA, ETHIOPIA

THE LEVEL OF ROUTINE HEALTH INFORMATION SYSTEM DATA QUALITY AND ASSOCIATED FACTORS AT PUBLIC HOSPITAL IN SILTE ZONE, CENTRAL REGION, ETHIOPIA, 2023: A MIXED STUDY

BY: NIGUSSIE DUKAMO (BSC.NURSING)

ADVISORS: 1. DR. MIHIRETU BELAYNE (PHD)

2. MR. AGUNE ASHOLIE (MSC, MPH)

A THESIS SUBMITTED TO HAWASSA UNIVERSITY COLLEGE OF MEDICINE AND HEALTH SCIENCE, SCHOOL OF PUBLIC HEALTH IN THE PARTIAL FULFILLMENT OF MASTERS PROGRAM IN EPIDEMIOLOGY.

NOVEMBER, 2023

HAWASSA, ETHIOPIA

Declaration

I declared that this MPH research is my original work and has not been presented for degree in this or another university and that all sources of material used for this research have been fully acknowledged.

Name: Nigussie Dukamo (BSc)

Signature:_____

Date: _____

ADVISORS APPROVAL SHEET

This is to certify that the thesis entitled “the level of routine health information system data quality and associated factors at public hospital in Silte zone, Central region, Ethiopia, 2023 in partial fulfillment of the requirements for the degree of Master’s with specialization in Epidemiology, the Graduate Program of the School of Public health and has been carried out by Nigussie Dukamo Id. No GPPHEPW/0012/13, under our supervision. Therefore we recommend that the student has fulfilled the requirements and hence here by can submit the thesis to the department.

Name of major advisor	Signature	Date
Mihiretu Belayneh (PhD)	_____	_____

Name of co-advisor	Signature	Date
Agunie Asholie (MSc, MPH)	_____	_____

EXAMINERS APPROVAL SHEET

SCHOOL OF GRADUATE STUDIES HAWASSA UNIVERSITY EXAMINERS' APPROVAL SHEET

We, the undersigned, members of the Board of Examiners of the final open defense by have read and evaluated his thesis Entitled "The level of routine health information system data quality and associated factors at public hospital in Silte zone, Central region, Ethiopia, 2023", and examined the candidate. This is, therefore, to certify that the thesis has been accepted in partial fulfillment of the requirements for the degree.

_____	_____	_____
Name of the Chairperson	Signature	Date
_____	_____	_____
Name of Major Advisor	Signature	Date
_____	_____	_____
Name of Internal Examiner	Signature	Date
_____	_____	_____
Name of External examiner	Signature	Date
_____	_____	_____
SGS Approval	Signature	Date

Acknowledgment

First, I would like to thank you the Hawassa University College of Medicine and Health Sciences for approval of the ethical concerns to conduct this research.

I would like to express my deepest gratitude to my Advisor; Dr Mihiretu Belayneh (PhD), who helped me a lot in guiding the way and for his constructive ideas; he gave me their unreserved support and guidance to develop this research paper.

My acknowledgment also goes to DUP/DDCF (Doris Duke Charitable Foundation) project for providing financial support to the success of this thesis report.

Lastly, my acknowledgment goes to the Silte zone Health bureau; Hospital higher officials as well as study participants and data collectors for their valuable involvement and contributions.

Table of contents

List of Table

List of figure

Acronyms and abbreviations

AA Addis Ababa

ART Antiretroviral therapy

AOR Adjusted Odd Ratio

BSc Bachelor of Science

CI Confidence Interval

CBMP Capacity Building and Mentorship Project

DDCF Doris Duke Charitable Foundation

DHIS District Health Information Software

DQR Data Quality Review

IRB Institutional Review Board

FMOH Federal Ministry of Health

GDDS General Data Dissemination Strategy

HIS Health Information System

HMIS Health Management Information System

HSTP Health Sector Transformation Plan

IT Information Technology

ICT Information and Communication Technology

IMF International Monetary Fund (IMF)

KII Key Informant Interview

LMICs Low- and Middle-Income Countries

LQAS Lot Quality Assurance Sampling

MOH Minister of Health

NCoD National Classification of Disease

NGO Non-governmental Organization

PCV Pneumococcal vaccine

PRISM Performance of Routine Information System Management

PH Primary Hospital

PMT Performance Monitoring Team

RHIS Routine Health Information System

SNNPR South Nations, Nationalities and People Region

SPSS Statistical Package for the Social science

TB Tuberculosis

VF Verification Factor

WHO World Health Organization

WCSH Werabe Comprehensive Specialized Hospital

Abstract

Background: *Routinely collected data of poor quality can compromise the validity of effect evaluations and lead to poor decision-making, inappropriate resource allocation, and a loss of trust in the health system. Routine health information system data are seen as poor in quality andare not used for decisions in Ethiopia, and continues to be a significant problem.*

Objectives: *To assess the level and associated factors of routine health information system data quality; and explore the factors affecting data quality at public hospitals in Silte Zone, Central Region Ethiopia, 2023*

Methods and materials: *Facility based mixed method with an embedded study design was conducted. A total of four public hospitals, 32 units or departments, 605 health care workers, and12 key informant interviews were selected. Simple random sampling and purposive sampling techniques were used for selecting study participants in quantitative and qualitative study. The data was entered into Epi-data version 4.4, open code 4.03 and exported to SPSS version 26, and descriptive statistics were used to assess the level of data quality, and Binary logistic regression and thematicanalysis was run to identify factors affecting data quality. Adjusted odds ratios with 95%confidence intervals and themes or subthemes were reported.*

Result: *The overall study conducted facility data quality level was 90.84%; completeness and consistence of data in this study were 85.5% and 85.3% respectively. Easily understandability of registration and report format [AOR 1.92; CI 1.11-3.33], receive training [AOR 1.62; CI 1.07- 2.44], getting supervision [AOR 1.66 CI 1.05-2.61], provide regular feedback [AOR 1.72 CI 1.07-2.75], team's work appreciated and valued by supervisors [AOR 1.61; CI 1.04-2.75] and decisionsand follow up actions identified in Performance monitoring team meetings [AOR 1.73; CI 1.12- 2.67] were significantly associated with data quality ; and thematic analysis was done and categorized into four themes and twelve sub-themes.*

Conclusion and recommendation: *The level of data quality at the public hospital in Silte Zone is almost equal to the national expected level of data quality, but completeness &consistency of the data were lower than the national expected level.*

For MOH and other supporting organization to fully transforming the paper based service registration to an electronic-based medical recording system in which it will reduce incompleteness and inconsistency of data and to increase data quality

Key word: *Data quality, Routine Health information System, Accuracy, Completeness, Timeliness, Consistency, Central region, Ethiopia*

1. Introduction

1.1 Background

According to WHO definition Health Information System (HIS) as a system that integrates data collection, processing, reporting, and use of the information necessary for improving health service effectiveness and efficiency through better management at all levels of health system(1).

A routine health information system (RHIS) is a system designed to gather, process, use, and dissemination of health-related data to enhance the management of programs, resources, and health care outcomes (2). Information from many parts of the health system is gathered via a routine health information system (3)

It have been used globally for more than a century. Recently, the developing world has begun to place a greater emphasis on RHIS, and it has becoming more well-known in the developing world. In 2008, Ethiopia began using the Health Management Information System (HMIS), which is meant to generate routine data for decision-making at various levels of the health system (4).

The information revolution is one of the big agendas of Ethiopia's Health sector transformation plan II (HSTP-II) and it is the phenomenal advancement in the methods and practice of collecting, analyzing, presenting, and disseminating information (5).

Data quality, defined as data's fitness to serve its purpose in a given context in terms of accuracy, completeness, consistence and timeliness is an essential element of this information revolution agenda (5). The quality of routine health information system data is vital for the health system to function well and for policymakers to be able to evaluate the effects of health system efforts to improve the health of the population (6).Routine health care data have no importance unless it is accurate, processed, and used for decision making hence responsive to the local situations (7). Improved health system performance is directly linked with the quality and use of routine data in a country's HIS (8). Despite the improvements in the knowledge and understanding of the role of health information in the global health system, the quality of data generated by routine HIS is still very poor in low and middle-income countries (9).

The International Monetary Fund (IMF) has developed General Data Dissemination Strategy (GDDS) to help countries in improving the quality of their data, the GDDS strategy provides a framework for evaluating needs for data improvement and setting priorities; it provides guidance on dissemination to the public of comprehensive, timely, accessible and reliable economic, financial and socio demographic statistics (10)

All functions of the health system and public health policy are seriously reliant on the presence and use of quality HIS data (11). However, lack of quality data and poor usage are affecting the health system's performance and the health of the society. This is evident by frequent over and under stocks of supplies, poor detection and management of outbreaks, and scarcity of human resources at different times (12).

Therefore, this study was assessed Level of routine health information system data quality and associated factors at public hospital in Silte Zone, Central region Ethiopia.

1.2 Statement of problem

HIS is the primary information system under Ethiopia's "One Plan, One Budget, and One Report"policy.One of the big agendas of Ethiopia's Health Sector Transformation Plan II (HSTPII) is the information revolution, which is the phenomenal advancement in the methods and practice of collecting, analyzing, presenting, and disseminating information. The data quality elements are essential for the information revolution agenda (13).

Health systems in low- and middle-income countries are still suffering from a suboptimal quality and inadequate use of data generated by their routine health information systems (14, 15).

The performance of routine health information systems (PRISM) framework identifies the potential influences on the performance of routine health information systems, which are categorized as technical, behavioral, and organizational influences (16).

Incomplete registers, a lack of consistency between registers and reports, and a low level of data accuracy are common types of data quality problems (17-19). The quality of data was found to be between 34 and 72% in many African countries (20)

Studies were done on routine health information systems data quality; in Oyo State, Nigeria stated that, an average completeness (89.42%) was found in the monthly summary form, 65.24% in the DHIS computer software, and an overall average of 77.33% (21) with verification factors of 1.16 for all sampled health facilities (21). The data from Benin, showed that the health workers competence associated with their training and their work engagement as a sign of motivation, was associated with the quality of the data, and the findings support those in Uganda, which highlighted the relationship between behavior and RHIS performance, particularly regarding the use of the information (22, 23).

The study was carried out in Ethiopia at South nations, nationalities and people region (SNNP), where the data items totally matched between the registers and the reports in roughly 41%, less than 7% of the facilities failed to submit their monthly forms in a timely manner, and five out of eight Woreda health offices in the Kembata Tembaro zone lacked the necessary records. Records of report receipt at the fourteen Woreda health offices revealed that 77% of the healthcare facilities met the reporting deadline (24), another study at Harari region, Ethiopia stated that, only 58.1% of departments had accurate data, and the lowest proportion of accuracy 45.6% was observed in the health posts, and overall 40% of the departments have incomplete data where the majority

93.7% of the study units submitted their report on timely. Study done at Harari region, 51.35% departments have good data quality moreover, more than one third 38.83%, about two third 65.85% and more than half 54.05% of the departments at the health posts, health centers and hospitals respectively were found to have good quality data (13).

The study data in East Wollega showed that the registration content completeness was 78.2% whereas, report content completeness was less than the 86%, indicating that the health workers focus more on managing patients than recording data due to the work load and lack of commitment to the data(25)

The data accuracy may be affected by errors that occur during data entry, intentionally manipulating the data for different reasons possibly competition among the staffs and facilities, false reports to increase achievement, reports not made on time, lack of trained personnel able to fill out the reporting formats, feedback, and type of facility were significantly associated with the data quality (13)

Improved data quality leads to better decision-making across an organization so, excellence in data quality enables health care organizations to plan and provide effective and efficient service for users.

Even though supportive supervision, short-term training, and the data owners have been assigned, routine health system data quality is still a major problem in low- and middle-income countries, including Ethiopia. This is affecting the health system performance and the health of the society, as evident by the frequent over- and under stocking of supplies, the poor detection and management of outbreaks, and the scarcity of human resources at different times (12).

The core determinants of routine health information system data quality and utilization are technical, behavioral and organizational factors (26, 27). However, less is known about evidence-based actions to successively progress the use of information for decision-making to improve data quality, effectiveness, and efficiency of health service delivery despite significant attention given by the Ministry of Health of Ethiopia to the routine health information data quality improvement.

Few studies have been done in Ethiopia on the routine health information system data quality by combining data from health posts, health centers, health offices, and hospitals at the same time. These studies also included few program-specific indicators and used only three dimensions to

measure data quality. However, from what I have read, there is no evidence showing the level of routine health information system data quality and associated factors at public hospitals in the Silte Zone. For this reason, in addition to the previously mentioned gap, this study attempted to address the data quality at public hospital level, included various program areas, and applying four dimensions to assess the data quality at each program area or unit.

1.3 Significance of study

The result of this study was primarily assist decision makers to understand the existing level of routine health information system data quality in the area.

Also, the identified findings and the recommendations forwarded to regional and Zonal health bureau and partners who are currently working in collaboration with local decision makers to tackle identified gaps.

Furthermore, the study findings are the input for the national level policy makers (FMOH and other stakeholders) to know and evaluate the status of current national plan which is information revolution agenda.

Additionally, it adds value to the current literature and could be used as a source of information for future studies of a similar nature.

2. Literature review

The health information system (HIS) is one of the six building blocks of a health system designed for the generation and use of information for other functions of the health system (28).

The purpose of a health information system is to routinely generate quality health data that provides specific evidence support to make decisions on health issues (29). The information revolution is one of the big agendas of Ethiopia's Health sector transformation plan II (HSTP-II) and it is the phenomenal advancement in the methods and practice of collecting, analyzing, presenting, and disseminating information (5).

All functions of the health system and public health policy are seriously reliant on the presence and use of quality HIS data however, lack of quality data and poor usage are affecting the health system's performance and the health of the society (12, 30, 31).

The Performance of Routine Information System Management (PRISM) framework categorized the factors that influence the data quality in to three groups: organizational, technical, and behavioral factors (27). The variables which are related to behavioral factors which affect data quality are level of knowledge, negligence, and data manipulation for competitive reasons, motivation, and sense of responsibility (32, 33). Technical factors affecting the quality of the data include the user friendliness of reporting formats, trained person able to fill format and standardized indicators, while organizational factors include the availability of training, feedback, supervision, and data utilization (34, 35).

The following section review different studies result based on variables which have potentially associated with routine health information system data quality.

2.1 .The quality of routine health information system data

The RHIS has been proven to be a very useful tool for monitoring and making adjustments to resource allocation and policy implementation (36). The Ethiopian Federal Ministry of Health has responded by reforming and redesigning the national RHIS, the reform has taken significant measures to address a lack of routine health data, which has hindered the quality of care, planning, and management systems, as well as decision-making (37)

The quality of routine RHIS data in low- and middle-income countries remains quite low in the global health system (9). In India, Nepal, and Pakistan, studies show that the overall health data quality was much below the national standard (38). In many African countries, data quality was found to be in the range of 34%–72% (24).

Ethiopia's routine health information system found that data quality is below the national average of 90%, In study that was conducted at Dire Dawa, the overall data quality in the unit or department was found to be 75.3% (39), Previous evidence in Ethiopia, including the South Nation Nationality People Region, suggests that the level of data quality was recorded as below the national threshold (40)

2.2. Dimensions of data quality

Data quality referred to the "fitness for use" (41), More recently, researchers propose that there are properties of data that determines data quality and there appears to be an agreement that data quality is a multidimensional concept (42), However, though there is no agreement on the dimensions of data quality there are cross cutting dimensions identified by the literature: completeness, timeliness, consistency, accuracy, reliability and precision (23). Specifically, the dimensions; completeness, accuracy, consistency and timeliness were found the most common reviewed in the literature (42).

According to a study on the assessment of data quality in routine health information systems in Oyo State, Nigeria, the monthly summary form's average completeness was 89.42%, the DHIS computer software's average completeness was 65.24%, and the overall average was 77.33% (21), and the Mozambique study found that the indicators they used for their study had an accuracy range of 91-97% and a completeness of over 90% (25).

The study conducted in Ethiopia at SNNPR stated that overall, about 41% of cases, the data items completely matched between the registers and the reports. Only at the hospital level, the data accuracy for all the data elements exceeded the minimum level of 60% accuracy, and the completeness of the monthly report is measured by the number of health facility reports with over 90% of the data elements filled against the total number of data elements that the facility was supposed to fill, but the result of the study showed that less than 7% (93% of health facility reports meet 90% completeness criteria) of the facilities did not complete the monthly form before

reporting, and five out of eight Woreda health offices in Kembata Tembaro zone did not have records to measure timeliness. The fourteen Woreda health offices had records of report receipt that showed 77% of the health facilities met the reporting deadline (24), but according to a study done by Addis Ababa, the overall regional data quality was 76.22%, ranging from 68% at Yeka Health Center to 92% at Shiromeda Health Center. Out of a total of 4752 monthly reports, 148 were outliers. Thus, around 3% of the reports from the city of Addis Ababa health center were inconsistent by comparing the data from Penta3 and PCV3, consistency between related metrics is assessed (2).

In the health centers run by the Addis Ababa city administration, the average accuracy report is 77.67%, with a standard deviation of 9.65. The accuracy of the monthly reported data from only five (15.5%) health centers fall within the acceptable range of accuracy (90%–110%). The timeliness of the health centers report was evaluated using the DHIS2-generated timelines report; the health centers median score for report timeliness ranged from 0% to 100%, while the percentages for report and source document completion were 93.93% and 96%, respectively (2).

Another study done in Ethiopia at the Harari region found that, of the 222 departments whose data accuracy was examined, 129 (58.1%) had accurate data, with the health posts having the lowest accuracy percentage (45.6%). From the 5230 cases checked for registration content completeness with the relevant information, more than two-thirds (69.6%) of the cases were completely registered on the registration. Overall, this study revealed that about 89 (40%) of the departments have incomplete data, whereas the rest, 133 (60%) have complete data, and the majority (93.7%) of the study units submitted their report on time, while only 14 (6.3%) did not (13).

2.3 . Organizational factors that affect the quality of RHIS data

These factors can be the type and size of facility, culture, politics, hierarchy, planning and control system, strategy, management, and communication. The PRISM framework considers organizational determinants key for affecting performance. It defines this category as all those factors that are related to organizational structure, resources, procedures, support services, and culture to develop, manage, and improve RHIS processes and performance (27).

In addition to organizational structures, such as the availability of sufficient room for HMIS activity, external factors like inadequate supporting infrastructures, like electric power supply, poor road transportation, and telecommunication affect RHIS performance significantly (43)

The study conducted in Uganda reported a prevalence was 55% for poor data quality (22), in Benin stated that, availability of resources, supervision, financial incentives, and the perceived complexity of the technical factors were not associated with the quality of the data (23).Study conducted at Addis Ababa city administration in Ethiopia revealed that a lack of supplies such as registration books, tally sheets, and infrequent and inconsistent supportive supervision and workload, were contributed to poor data quality, for this study key informants mentioned that a lower motivation level as a root cause for poor data quality (44)**,** from this study, the non-functional PMT meeting and inefficient training for the skill gap despite the effort as the main cause of the data quality problems according to key informant interview (45) .

A study in Somali Regional State, eastern Ethiopia, found that health workers who felt their immediate supervisors or PMT did not value their data recording were less likely to ensure data quality (46). Study in Harari Region, Ethiopia, revealed that (77.5%) of participants reported to have received supervision and (14.9%) of them have received refreshment trainings on HIS in the last 6 months, and 61.7% of the departments have received the feedback (13).

At Hadiya Zone, those supervised healthcare workers were 3.5 times more likely to report quality data compared to those who were not supervised; those who completed the registration were 2.7 times more likely to report good quality data than those who did not; and those who check data quality in the health facility were 1.3 times more likely to report quality data compared to those who were not (47).

2.4. Technical factors that affect the quality of RHIS data

Technical determinants are factors related to the friendliness of the reporting format, using a standardized format, and the technology required to develop, manage, and improve the quality of RHIS data and its performance. It also comprises information technology types as well as software development for data processing and analysis (48). The effect of technical factors on RHIS is supported by an empirical investigation on data warehouse adaptation; the study claims the complexity of IT infrastructures is a key determinant for the adaptation of new information systems

(49). Besides having the right user attitude and skills with good leadership, designing a user-friendly health information system is inevitable for data quality improvement (50).

Study conducted in India Concerned about poor data quality stated that, simplifying formats, reducing redundancies, minimizing duplication, and utilizing ICT to encourage health systems to create and use data (51).

A study in the Harari region of Ethiopia revealed that the majority of the departments 183 (82.4%) have standardized indicators, 178 (80.2%) said their reporting forms are user-friendly, 174 (78.4%) have trained staff who can fill out the reporting formats (13). In another study on data quality at health centers in southern Ethiopia's Hadiya Zone, 59.8% of respondents said their departments used a set of standard indicators, including case definitions. A huge 77.7% of respondents said that HMIS is an approachable format, and 40.5% said that there is qualified staff available to compile data and fill out forms (40).

2.5. Behavioral factors that affect the quality of RHIS data

In addition to technical and organizational factors, individual-level factors affect the practice of RHIS tasks (52, 53). If people appreciate the usefulness of RHIS tasks, feel confident and competent in performing the task, and perceive that the task's complexity is challenging but not overwhelming, then they will complete the task persistently (27).

In the study at Benin, the competence of the health workers was related to their training and their work engagement as a measure of motivation, which were related to the quality of the data (23).

More than two-thirds (69.4%) of the 216 department heads in the Harari region who participated in the study had a positive attitude toward HIS operations. In addition, just 48 (21.6%) had a good understanding of the rationale behind routine HIS data, and around 87 (39.2%) reported that data manipulation was present in their departments. Nearly one third (39.2%), or 70, reported that neglect was present (13); another study from North Gondar, Ethiopia, also reported that only 23.8% of staff received HMIS-related training (29). Whereas a study conducted in Addis Ababa reported that 49.5% of staff and 98.4% of HMIS focal persons are trained, findings from this study also support this argument, where 52.72% of health professionals did not consider data recording as their duty (2).

According to research conducted in the Hadiya Zone of Ethiopia, 59.8% of respondents said that their departments used a standard set of indicators, including case definitions, and 40.5% said that they had skilled staff members who could compile data and fill out forms. An average confidence

level of respondents was 63%, and an average perception (motivation) of people toward HMIS use and meaning was 49.1%. About 28% of respondents in this study reported the availability of incentives for HMIS activity, which is a training opportunity, and about 60.8% reported having knowledge of HMIS and about 66% reported having skill with data quality checking (40). A mixed study conducted in Addis Ababa revealed that key informants, when asked about the overall task competence of data clerks and health professionals about data quality, replied that it goes to the extent of reporting wrong data elements (45).

2.6 . Conceptual frame work

The conceptual framework was adapted from the PRISM framework and reviewing published literature (Ref no; (21, 27, 48, 49)).According to the framework and review, the following four major factors (socio demographic factors, organizational factors, behavioral factors, and technical factors) were affecting routine health information system data quality.

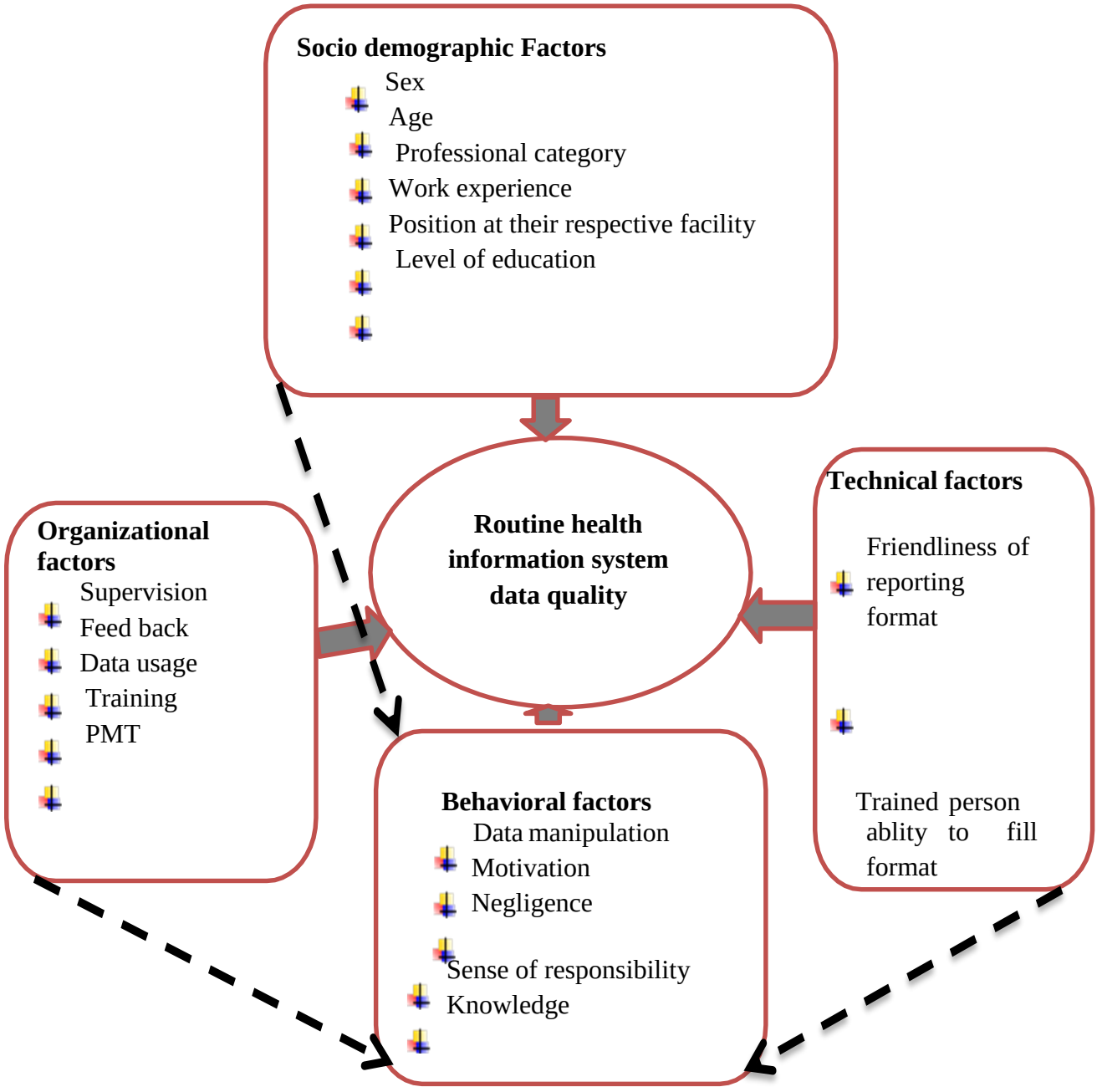

Figure 1:- Conceptual framework of factors which affect routine health information system data quality.

3. Objectives of the study

3.1. General objective

To assess the level and associated factors of routine health information system data quality; and explore the factors affecting data quality at public hospitals in Silte Zone, Central Region Ethiopia, 2023

3.2. Specific objectives

To assess the level of quality of routine health information system data at public hospitals, 2023

To identify factors affecting the quality of the routine health information system data at public hospitals, 2023

To explore the factors which affecting the data quality of the routine health information system at public hospitals, 2023

4. Method and materials

4.1. Study area

This study was conducted in Central region of Ethiopia, Silte zone at public hospital. Silte zone is found in Central region of Ethiopia and is 173 K.M away from Addis Ababa, capital city of Ethiopia. The zone has 212 kebeles, 12 districts and 3 administrative towns. According to 2018 CSA reports the average estimated population of zone is 1,017,557. Currently Silte Zone has four hospitals (Werabe comprehensive specialized hospital, Kibet Primary hospital, Tora primary hospital, Alem gebeya primary hospital) and one private medium clinic and totally there are 1184 health care workers with different disciplines in Silte zone public hospitals.(13)

4.2. Study design and period

A facility-based mixed method with embedded study design was employed at public hospital in Silte zone Central region of Ethiopia, from February 1 to March 30, 2023. A cross-sectional study design was used for the quantitative part and phenomenology study design was used for qualitative part.

4.3. Population

4.3.1. Source population

All public hospitals in Silte zone were taken as source population.

4.3.2. Study Population

The selected unit or department at public hospital in Silte zone were included.

4.3.3. Study unit

For quantitative study

HMIS registration books or report format as well as individual medical records and health care workers chosen randomly from particular departments or selected units.

For qualitative study

Key informants selected purposeful at public hospital

4.4. Eligibility criteria

4.4.1 Inclusion criteria

For quantitative study

Health care workers who have worked for more than six months at this institution were included in this study.

For three consecutive months, selected data that were registered in 2023 on the HMIS registration book were included

For quantitative study

The key informants, like HMIS officers, quality office members, and PMT members, were included in this study.

4.4.1. Exclusion Criteria

Health care workers who have worked for less than six months at this institution were excluded from this study.

4.5. Sample Size Determination

4.5.1. Sample size for objective one

WHO recommended for a regular data quality review to use five units with core indicators (54), and additionally, the countries may select other indicators based on their needs and the resources available. Based on this, eight units or departments at each hospital were included; there are four public hospitals, so the total sample size used to assess the level of data quality was 32 units.

4.5.2. Sample size for objective two

Sample size for objective two is calculated via Epi –info from the findings of previous similar studies. Calculation is separately made for three potential determinants (Training, Supervision and Feedback) which were consistently significant in many studies. To determine appropriate sample size, percent of outcome from unexposed group, two sided confidence level, power of study and relatively least extreme odds ratio to be detected. Based on this, the maximum calculated sample size is 605.

Table 1:-Sample size calculation for objective two (Reference (13, 40))

S.N	Variable	CI	% of outcome from unexposed	AOR	Power	Total Sample	Add 5% of non-response rate
1	**Training**	**95%**	**46%**	**2.47**	**80**	**201**	**211**
2	**Supervision**	**95%**	**57.3%**	**1.71**	**80**	**576**	**605**
3	**Feed back**	**95%**	**33%**	**3.08**	**80**	**131**	**138**

4.5.3. Sample size for objective three

The qualitative study 12 participants (3 from each hospital) were selected for key informant interview (KII).

4.6. Sampling technique and sampling procedure

The FMOH guideline of HMIS says that all Health care workers who are involved in HMIS activities, starting from the daily register of the source document to the final report should be included (30)

Health care workers for the self-administered questionnaire were selected using a simple random sampling technique.

Participants in the qualitative study were selected purposefully to explore the factors affecting data quality at public hospitals.

The three-month documents of registration or reports were reviewed to check accuracy, completeness, timeliness, and consistency of at each unit or department. The first month was selected randomly by lottery methods, and then three consecutive months registered data were included for chart or document review.

The selected units or departments were: maternal health (ANC1), laboratory (malaria tested case), immunization (Penta3/PCV3), ART (currently on ART), tuberculosis (new or relapsed TB case), inpatient department, emergency department, and outpatient department.

The sample size was allocated to each hospital and then to units or departments based on the proportional allocation formula (nj= n /N* Nj). Finally, the study participants were selected by simple random sampling method.

Study participants' selection for quantitative study (for self-administered questionnaire)

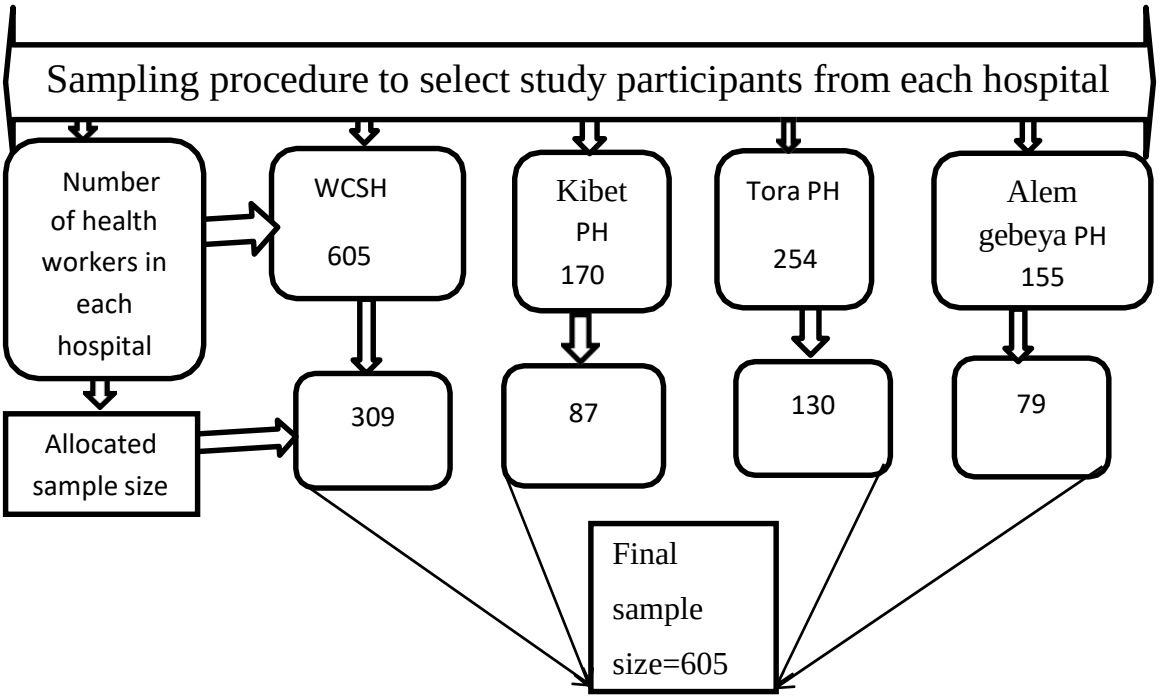

Figure 2:-Sampling procedure to get study participant from each hospital and finally sample size.

4.7. Variables

4.7.1. Dependent variables

Data quality (good or poor).

4.7.2. Independent variables

Socio-demographic related: sex, age, work experience, professional category, and level of education, salary and position in hospital.

Organizational related: feedback, supervision, training and data use, appreciation and performance monitoring meeting and accountability for poor performance.

Technical related: complexity of the reporting format or user friendly of formats and trained person ability to fill format.

Behavioral related: - Data manipulation for competition, negligence, sense of responsibility, and perception of staff on data, actively engagement and knowledge of routine health information data.

4.8. Data Collection Tool and procedure

For quantitative study

The data were collected using structured questionnaire, which is adapted from reviewing different literatures and PRISM tool (27, 49). Self-administered questionnaire contains four parts; socio demographic characteristics; technical factors; organizational factors and behavioral factors; and knowledge assessing parts.

Six data collectors (two Bsc nursing with HMIS training) and two supervisors with Master of Public Health were assigned.

These data collectors were collected the data from each respondents and review the document or registration as well as individual patient chart.

The data were collected by going to all the hospitals, explaining the aim of the study, ensuring the confidentiality of the data, and obtaining consent from each participant in each facility.

For qualitative study

A key informant interview is the method used to collect qualitative data. After a convenient time and place were selected for the key informants, a semi-structured interview guide was used.

Two Master of Public Health data collectors were used to collect qualitative data.

The data collectors conduct the interview using an audio recording device for 20–40 minutes while simultaneously taking field notes.

4.9. Data quality assurance

For quantitative study

The data collection tool was commented on by my advisers. An amendment of the tool was made after pretesting on 5% (30) of health care workers or 5% (2) charts or HMIS register of the total sample at Adare General Hospital, and the questionnaire was checked for clarity, simplicity, understandability, and consistency.

Before the data collection, two days of training were provided on the purpose, how to collect data, and ethical issues, emphasizing the importance of the safety of the participants and the quality of the data. Data collectors were supervised, and on-site technical assistance was given. Moreover, data completeness and consistency were evaluated on a daily basis, and corrective steps were implemented on time. Lastly, prior to data entry, each questionnaire was coded.

For qualitative study

The trustworthiness of the data was evaluated using the following criteria:

Credibility: During in depth interview, enough time given to participants to respond their perceptions and experiences, interviewed in comfortable place, data was collected by the supervisor and other data collector observe their facial expression to understand their nonverbal response.

Dependability: In order to assure the consistency of data from all KII was done by the same data collectors using the same KII guide. After data collection the raw or recorded data was transcribed verbatim and then translated into English language

Transferability: Nominated sample was used for in depth interview in order to have representative of the source population.

Conformability: In order to avoid researcher's bias throughout data collection, coding, and analysis, the KIs' own words were used instead of the researchers' opinions and expectation.

4.10. Data processing and Analysis

For quantitative study

The data was checked for completeness and consistency and then entered into Epi-data version 4.10.1.1 and exported to SPSS version 26 for further statistical analysis. Descriptive statistics mean,frequencies, and tables were used to summarize and describe the data. Mean scores are used as cut-off points to split the data into different scale measures to dichotomize variables. Binary logistic regression was performed, and the variables with a p-value < 0.25 were entered into the multivariable logistic regression analysis.

The Hosmer-Lemeshow statistical test was used to assess the model's goodness of fit (fitness if the p-value is greater than 0.05), the value of the variation inflation factor (VIF) was used to assess the Multicollinearity effect. If the mean VIF value of the independent variables in the regression model is less than 10, there is not an extreme collinearity problem.

Finally, the adjusted odds ratio (AOR) with its 95% CI was reported. Variables having a 95% confidence interval that does not include one in the multivariable logistic regression analysis were statistically significantly associated with the routine health information system data quality.

For the qualitative study

The data was collected from KII via audio recording and note-taking, then transcribed after listened repeatedly to cross-check the credibility and compatibility of the actual words of the respondent.

Translated verbatim from Amharic to English by Google Translate and then checked to see if it was coherently translated and data was saved in the form of plain text.

After reviewing and familiarization, responses from KIIs were coded, categorized, and analyzed using thematic analysis by open-code software, then themes and sub-themes were identified.

Finally, the themes were explained at the result and triangulated with the quantitative findings in the discussion in narratives.

4.11. Operational Definitions.

Technical determinants: - The factors that relate to the availability of HMIS tools, the complexity or user friendliness of formats, and the trained person's ability to fill out the format to perform routine health information tasks in certain organizations (55).

Organizational determinants: - Factors related to feedback, supervision, training, data use and performance monitoring team that contribute to improving the RHIS process and performance (55).

Performance monitoring team (PMT):- In this study, the performance monitoring team (PMT) are health care workers who are reviewing and analyzing the hospital's performance and develops action plans for course correction.

Behavioral determinants: - Individual-level factors affect the practice of Routine health information system tasks (52).

Data quality: - was measured by calculating the sum of the four dimensions of data quality measured and then taking the average of the scores. (56).

Good data quality: - The data that fits the criteria for the four quality dimensions, which means average scores of the four dimensions ≥ 90 (56, 57)

Poor data quality:-The data that does not fit the four criteria, which means average scores of four dimensions < 90% (13, 56, 58).

Completeness:- Is the average of the source document or registration content completeness and report content completeness, the data is complete if the average is $\geq 90\%$ (13, 40).

Data accuracy:-was measured by calculating the number from source document or register divided by the number from report submitted to the next level, the data is considered accurate if the average is between acceptable limit (0.90–1.10 or 90–110%), 10% tolerance for data accuracy was used (40, 58).

Report timeliness:-was measured by the number of reports delivered up to deadline for facility head divided by the number of reports expected to come, The data is timely if the average is $\geq 90\%$ (13).

Consistence: - was measured by the data on the register with data on individual medical records. By dividing those MRNs with matched data by the total number of sampled MRNs, the data is considered consistent if the consistency score is ≥ 90. (13)

Matched: - is defined as when all of the selected data elements that are recorded on the registers for the sampled individual are also recorded on the individual medical records.

Not Matched: - is defined as when at least one or more of the selected data elements that are recorded on the registers for the sampled individual are not exactly the same as what is recorded on the individual medical records. In addition, if the individual medical record is not physically available, then it is also considered not matched. (56).

Level of knowledge: - A health care worker is said to have good knowledge if they respond to knowledge questions above the respondent mean score.

4.12. Ethical Considerations

Ethical approval was obtained from Institutional Review Board (IRB) of Hawassa University College of medicine (Ref. No: IRB/184/15). The letter was submitted to Silte Zone health bureau then public hospital and official permission was written from them and submitted to the each unit or department head of hospital to get permissions of data collection. Study participants were informed about the purpose of the study, and informed consent was secured from study participants.

5. Result

5.1. Socio demographic characters of respondents

A total of 605 respondents were participated in this study with response rate of 100%. There are eight indicators in the units or departments that were included. From all facilities included in study, 575 (95%) were health care workers working at staff position, and most of the respondent's age was below 31, 389 (64.3%). Among the respondents, 313 (51.7%) were male. Regarding the distribution of levels of education, 474 (78.3%) were degree holders. About 258 (42.6%) of the respondents were nurses, and most of the respondents have years of experience below five year, in number 331 (54.7%).

Table 2: Socio-demographic characteristic of respondents at public hospital in Silte zone, Central region, Ethiopia, 2023 (n=605)

Variables	Category	Frequency	Percent
Age in year	<31	389	64.3
	>=31	216	35.7
Sex	Male	313	51.7
	Female	292	48.3
Year of service in year	<=5	331	54.7
	6-10	215	35.5
	>=11	59	9.8
Position	Health care workers (staff)	575	95.0
	Case team leader	26	4.3
	HMIS officer	4	0.7
Educational status	Diploma	116	19.2
	BSC degree	474	78.3
	Master's Degree	15	2.5
Salary in ETB	1046-4095	3	0.5
	4096-12695	601	99.3
	>12695	1	0.2

5.2 Level of data quality

The study's overall data quality was 90.84%, with a 95% confidence interval of [88.9–92.76]. The lowest data quality was found at Tora PH, while the highest percentage was found at Alem Gebeya Primary Hospital, at 89.83% and 93.2%, respectively.

Table 3: The level of data quality at each public hospital in Silte Zone, Central Region, Ethiopia, 2023 (n = 32)

Name of facility	Accuracy	Completeness	Timeliness	Consistence	Data quality
WCSH	98.71	79.54	100	81.95	90.05
TPH	97.53	85.18	91.67	84.95	89.83
Kibet PH	96.30	89.10	87.51	88.19	90.28
Alemgebeya PH	98.54	88.28	100	85.99	93.20
Over all	97.77	85.52	94.79	85.26	90.84

5.3 The level of data quality dimensions at each unit

5.3.1 Data quality dimensions at each unit in the Werabe comprehensive specialized hospital

In order to check accuracy of the data at WCSH, the top six most common diseases were selected, and three-month disease or service registers were counted from the HMIS registration book. A total of 3291 diseases or services were reviewed. Out of these, 0.89% (111/112) of the IPD data, 2.04% (576/588) of the emergence data, and 0.59% (1338/1346) of the OPD data were over reported. Three selected months at WCSH, 4920 clients were registered at various departments or units. Of those, 246 (or 5% of the register) data were checked for consistency and completeness at each department or unit.

The three-month report submission dates were checked at each department or unit and ensured that all reports were sent on time to the HMIS unit or higher level.

Table 4: The level of data quality dimensions at each units or departments at WCSH in Silte zone, Central region, Ethiopia, 2023 (n=8)

Units or departments	**Level of data quality dimensions**				
	Accuracy	Completeness	Timeliness	Consistence	DQ level
TB unit(New or relapsed case)	100	83.00	100	80.00	90.75
Laboratory unit (Malaria tested)	93.3	80.2	100	87.7	90.3
Maternity unit (ANC1)	100	85.3	100	76.5	90.45
Immunization (Penta 3/PCV3)	100	77.3	100	100	94.33
ART unit (New HIV)	100	80.4	100	100	95.10
IPD Department	99.1	88.9	100	72.2	90.05
Emergency Department	97.9	50.6	100	60	77.13
OPD Department	99.4	90.6	100	79.2	92.30
Overall result (%)	98.71	79.54	100	81.95	90.05

5.3.2 Data quality dimensions at each unit in the Tora Primary hospital

From the list of the ten most common diseases, the top six were selected, and the data accuracy at Tora Primary Hospital was checked by counting three-month disease or service registers from the HMIS registration book. A total of 2252 diseases or services were checked; of those, 1.95% (453/462) of emergence data, 2.86% (34/35) of TB data, 2.64 % (221/227) of OPD data, and 6.15% (61/65) of malaria data were over reported. 2420 clients were registered at various units or departments for three particular months; in which 121 (5% of total register) clients register were checked. One month emergence and OPD report were not submit timely.

Table 5: The level of data quality dimensions at each units or departments at TPH in Silte zone, Central region, Ethiopia, 2023 (n=8)

Units or departments	**Level of data quality dimensions**				
	Accuracy	Completeness	Timeliness	Consistence	DQ Level
TB unit (New or relapsed case)	97.1	89	100	80	91.53
Laboratory unit (Malaria tested)	93.8	83.5	100	86.4	90.93
Maternity unit (ANC1)	100	81.8	100	81.9	90.93
Immunization (Penta 3/PCV3)	100	83.6	100	91.9	93.88
ART unit (New HIV case)	100	99	100	100	99.75
IPD Department	100	80.4	100	83.4	90.95
Emergency Department	91.9	79.5	66.7	81.9	80.00
OPD Department	97.4	84.6	66.7	81.9	80.65
Overall result (%)	97.53	85.18	91.67	84.95	89.83

5.3.3 Data quality dimensions at each unit in the Kibet Primary hospital

In Kibet primary hospital, there were 1939 diseases or services counted and checked for data accuracy, of which 2.13% (46/47) of TB data, 5.66% (100/106) of IPD data, 3.45% (112/116) of Penta3 vaccine data, 3.82% (126/131) of OPD data, and 2.22% (309/316) of emergence data were over reported, whereas 7.48% (431/401) of ANC1 data and 0.12% (813/812) of malaria data were under reported. And, 2540 clients were registered at different selected units or departments, in which 127 (5% of the register) client registers were checked for data completeness and consistence. One-month laboratory (malaria) and emergence reports were not submitted timely.

Table 6: The level of data quality dimensions at each units or departments at KPH in Silte zone, Central region, Ethiopia, 2023 (n=8)

Units or departments		**Level of data quality dimensions**			
	Accuracy	Completeness	Timeliness	Consistence	DQ level
TB unit (New or relapsed case)	97.9	84.6	66.7	100	90.18
Laboratory (Malaria)	99.9	90.8	66.7	95.1	88.13
Maternity unit (ANC1)	92.5	86	100	82.2	90.18
Immunization (Penta 3/PCV3)	91.8	88.9	100	100	95.18
ART unit (New HIV)	100	100	100	100	100
IPD Department	94.3	86.8	100	80.0	90.28
Emergency Department	97.8	75.9	66.7	70.9	77.82
OPD Department	96.2	88.3	100	77.3	90.45
Overall result (%)	98.54	88.28	100	85.99	93.20

5.3.4 Data quality dimensions at each unit in Alem Gebeya primary hospital

From the service or disease register (2651) at three selected months in Alem Gebeya Primary Hospital, 2.22% (44/45) of IPD data and 0.19% (1072/1074) of malaria data were over reported, whereas 2.15% (285/279) of emergence data and 3.02% (648/629) of ANC1 service data were underreported in Alem Gebeya Primary Hospital, 3040 clients were registered at different selected units, and 152 (5% of the register) client registers were checked for completeness and consistency of data at each unit or department and one-month OPD reports were not submitted timely.

Table 7: The level of data quality dimensions at each units or departments at Alem Gebeya primary hospital in Silte zone, Central region, Ethiopia, 2023 (n=8)

Units or departments		**Level of data quality dimensions**			
	Accuracy	Completeness	Timeliness	Consistence	DQ Level
TB unit (New or relapsed case)	100	85.1	100	85.7	92.70
Laboratory unit (Malaria tested)	99.5	80.7	100	83	90.80
Maternity unit (ANC1)	96.9	90.1	100	78.1	91.28
Immunization (Penta 3/PCV3)	100	93.1	100	100	98.28
ART unit (New HIV case)	100	98	100	100	99.50
IPD Department	97.8	84.4	100	81.8	91.00
Emergency Department	97.9	86.5	100	82	91.60
OPD Department	96.2	88.3	100	77.3	90.45
Overall result (%)	98.54	88.28	100	85.99	93.20

5.4 Factors related to routine health information system data quality

The likert scale (five scales used), ranges from strongly disagree to strongly agree. These responses were dichotomized into disagree if answer were 1 to 3 and agree if answer were 4 and 5 codes.

Table 8.Organizational, behavioral and technical factors related to RHIS data quality at public hospital in Silte zone, Central region, Ethiopia, 2023 (n=605).

Variables	Categories	Frequency	Percent
Education	Diploma	116	19.2
	BSC	474	78.3
	Masters	15	2.5
Standardized set of indicators	Disagree	216	35.7
	Agree	472	64.3
Understandability of registration format	Disagree	133	22.0
	Agree	472	78.0
Trained staff able to fill out format	Disagree	250	41.3
	Agree	355	58.7
Received training	Disagree	264	43.6
	Agree	341	56.4
Getting supervision	Yes	405	66.9
	No	200	33.1
Providing regular feedback	Disagree	206	34.0
	Agree	399	66.0
Supervisors check data quality	Disagree	194	32.1
	Agree	411	67.9
My work is appreciated by supervisors	Disagree	245	40.5
	Agree	360	59.5
Staffs engage actively	Disagree	207	34.2
	Agree	398	65.8
Decisions are made in PMT meetings	Disagree	236	39.0
	Agree	369	61.0

Institution encourages	Disagree	233	38.5
	Agree	372	61.5
Staffs are accountable	Disagree	246	40.7
	Agree	359	59.3
knowledge of RHIS	Poor	285	47.1
	Good	320	52.9

5.5 Bivariate and multivariable analysis

In the bivariate logistic regression analysis educational status, knowledge of RHIS, presence of standardized set of indicators, registration and report format easily understandability, trained staff ability, receiving training (HMIS related activities), get supervision from higher officials, provide regular feedback to their staff, check data quality, collecting data is appreciated and valued by supervisors, staff engagement actively in all activities, decisions and follow up actions identified in PMT meetings, institution encourages to gather data, and accountable for poor performance had a p-value of less than 0.25 and included in multivariable model. In the multivariable logistic regression, the registration and reportformat is user-friendly or easily understandable; receiving training on HMIS-related activities; getting supervision from higher officials; providing regular feedback to their staff; collecting datais appreciated and valued by supervisors; decisions and follow-up actions identified in PMT meetings were significantly associated with routine health information system data quality.

Participants concurred that having registration and report formats that were easy to use or understand increased the likelihood of having good data quality by 1.9 times (AOR = 1.92; 95% CI: 1.11-3.33; P = 0.020) compared to others.

Healthcare workers who received HMIS-related training had 1.6 times increased likelihood of having good quality data compared to those who did not (AOR = 1.6; 95% CI: 1.07–2.44 with P = 0.022); and these who received supervision from higher-ranking officials had a 1.7 fold increased likelihood of having high-quality data compared to those who did not receive supervision (AOR = 1.66; 95% CI: 1.05–2.61 with P = 0.029); and those who received regular feedback from supervisors through reports had a 1.7-fold higher likelihood of having good quality data compared to those who did not receive regular feedback from supervision (AOR = 1.72; 95% CI: 1.07-2.75 withP = 0.024).

Health care workers who agreed on the impact of decisions made and follow-up actions identified in PMT meetings based on presented data on data quality had 1.7 times higher odds of having good data quality than others (AOR = 1.73; 95% CI: 1.12-2.67, P=0.013) and; in addition those whose work is appreciated by supervisors and coworkers have a 1.6 times higher odds of havinggood data quality than others (AOR=1.61; 95% CI: 1.04-2.48 with P=0.031).

Table 9**:** Bivariate and multivariable logistic regression result at public hospital in Silte zone, Central region, Ethiopia, 2023 (n=605).

Variables	Categories	Data quality		COR(95% CI)	AOR (95%CI)
		Good n (%)	**Poor n (%)**		
Education	Diploma	95 (81.9)	21 (18.1)	1	1
	BSC	347 (73.2)	127 (26.8)	1.7 (0.99,2.77)	1.69 (0.99,2.92)
	Masters	11 (73.3)	4 (26.7)	1.6 (0.48,5.67)	2.33 (0.60,9.01)
Presence indicators	Disagree	178 (82.4)	38 (17.6)	1	1
	Agree	275(70.7)	114(29.3)	1.9 (1.28,2.93)	1.36(0.87,2.12)
Registration understandable	Disagree	113 (84.9)	20 (15.1)	1	**1**
	Agree	340 (72)	132 (28)	2.19(1.31,3.68)	**1.92(1.11,3.33)***
Staff ability	Disagree	190 (76)	60 (24)	1	1
	Agree	263 (74.1)	92 (25.9)	1.1(0.76,1.61)	0.99(0.66,1.49)
Received training	Yes	239 (70.1)	102 (29.9)	1.83(1.24,2.68)	**1.62(1.07,2.44)***
	No	214 (81.1)	50(18.9)	1	1
Supervision	Yes	287 (70.9)	118 (29.1)	2.0(1.30,3.08)	**1.66(1.05,2.61)***
	No	166(83)	34(17)	1	**1**
Regular feedback	Disagree	173 (83.9)	33 (16.1)	1	1
	Agree	280 (70.2)	119(29.8)	2.2(1.45,3.42)	**1.72(1.07,2.75)***
Supervisor check	Disagree	154 (79.4)	40(20.6)	1	1
	Agree	299(72.8)	112(27.2)	1.4(0.96,2.17)	1.1(0.72,1.78)
Appreciation	Disagree	202 (82.4)	43 (17.6)	1	1
	Agree	251 (69.7)	109(30.3)	2.0(1.37,3.04)	**1.61 (1.05,2.48)***
Staff engagement	Disagree	165 (79.7)	42(20.3)	1	1
	Agree	288 (72.4)	110(27.6)	1.5(1.00,2.25)	1.21(0.78,1.88)

Decisions in PMT meetings	Disagree	194 (82.2)	42(17.8)	1	1
	Agree	259 (70.2)	110 (29.8)	1.97(1.31,2.93)	**1.73(1.12,2.67)***
Staff encourages	Disagree	184 (78.9)	49 (21.1)	1	1
	Agree	269(72.3)	103(27.7)	1.44(0.98,2.12)	1.04(0.68,1.59)
Accountability	Disagree	191 (77.6)	55 (22.4)	1	1
	Agree	262 (72.9)	97 (27.1)	1.29(0.88,1.88)	1.13(0.75,1.71)
Knowledge	Poor	222 (77.9)	63(22.1)	1	1
	Good	231 (72.2)	89 (27.8)	1.36(0.94,1.97)	1.19(0.80,1.77)

COR = Crude odds ratio; CI = Confidence interval; AOR = Adjusted odds ratio

*Statistically significant

5.5 Thematic Findings

Thematic analysis was done using open-code software and categorized into four themes, and twelve sub-themes. These themes were data collection tools and its impact, data quality challenges and assurance mechanisms, supervision and feedback on data quality, and training and its impact on data quality.

Theme one: data quality challenge and assurance mechanism

1. Challenge of data quality

A majority of in-depth interviewees mentioned that there was incomplete information on the register, the hospital does not use the HMIS data, negligence, or carelessness, and improper disease classification are major challenges that affect data quality.

The key informants were responded for the reason of incompleteness of information as follow:

Most of the time some registered data are incomplete. Because there are some careless or negligent health professionals, data will not be complete. In order to register indicators correctly and get quality data, incomplete registration is the main problem of our data quality. **[KII 2, HMIS officer]**

*Many times there is something that makes it difficult for us to maintain data quality, some health professionals do not register information, not one person who is trained in health management information system, but also all the health professionals who are assigned to do it in turn. It's a problem, sometimes, we may write a diagnosis that is not correct, and another register may be written in illegible handwriting. The other is incompleteness, this one is like another challenge, the registration book should have complete things, such as writing one by one, but some staff may be fill up by jumping. Many times, when trainings are given at the institutional level to solve this problem, we are told to provide complete information, few of them are filled with complete information, but when we visit the unit, what the situation we seen is incomplete card or register***.** **[KII 9, Quality directorate]**

The majority of individuals pointed out the challenge related to improper disease classification as follow:

Regarding the disease registration, many clinician expected write HMIS disease code, but with us, the doctor only writes a clinical diagnosis, then the data owner records those diseases from the registration, when data owner makes monthly report, to match it with National Classification of Disease (NCoD), they may not find the correct clinical diagnosis, so they rounds it off. **[KII 2, HMIS officer]**

The key informant responded to the challenge related to hospital does not use the HMIS data based on gaps and negligence or carelessness as follows:
As a hospital, there are some gaps, instead of using data, if we use the data to improve our services it is very good, but in reality, hospital did not use according to our plan. **[KII 12, Quality officer]**

I don't think that negligence has anything positive effect, because it affects the data quality very much. There are some professionals who do not register what they have done, some who start and make incomplete registers, for example, some professionals who are in dire need of emergency, do not record the time of patient arrival. Because of that, sometimes it is difficult to know patients who have been there for more than 24 hours. Because of this, negligence affects the data quality very much. **[KII 2, HMIS officer]**

Being negligent, for example, when Health professional is negligent, while filling out information they may changing for example gender to a man who is a woman might be registered and also HMIS diagnosis may or may not have been recorded correctly and this will affect report. At this time, the health professional does not know that not filling in the correct information will cause a gap in the quality of the information. Negligence has a big impact on data quality. **[KII 8, Planning &police director office]**

2. Assurance mechanism of data quality

A majority of in-depth interviewees mentioned that there was an assigned HMIS focal person, a strong PMT monthly meeting, and the LQAS (lot quality assurance sampling) method to assure data quality.

One participant explained that one of the data quality assurance mechanisms was checked by using the LQAS method, as follows:

When we come to data quality assurance methods, there is a registration book, there is a tally sheet and there is a monthly report. We will check with LQAS if there are selected indicators on all three of them. If they are under LQAS, they will work again. If it is correct, it will be sent to HMIS and then DHSI 2 will be entered. Quality assurance mechanism in our institution is majorly LQAS method. **[KII 4, PMT member]**

The key informant responded that one of the data quality assurance mechanisms was assigning an HMIS focal person and strengthening PMT monthly meetings

There was a matter of not registering the patient as soon as it came and not registering it due to different reason but now they have been assigned an independent focal person who follow it and also now whether it is elective or emergency case, there is assigned nurse to register. **[KII 1, Dep coordinator]**

There are performance monitoring team (PMT), so PMT report periodically, monthly data, weekly report, data we report at case team level, especially those monthly report data is evaluated at case team level by PMT, before being sent to our hospital HMIS department. After they are confirmed, or after the quality of the data items are confirmed, they are sent to HMIS department this is one of the ways to assure data quality. **[KII 7, Quality officer]**

Theme two: data collection tools and its impact

A majority of in-depth interviewees mentioned that tool availability and friendliness affect RHIS data quality, as stated as follows:

The tool itself will orient you like any register tool, tally sheet, report format, are easy to understand. If the Ministry of Health of Ethiopia's format is finished, we will print it in our institution. I don't think there is a problem with the data register tool because the tool is easy to understand, if you need an explanation, you will find it written below in the registration or report format. Tool availability and easy understandability has a positive effect. Currently we have a tool that you can use to do all staff to reports and registers. And that has a positive effect on data quality. **[KII 10, Nursing& midwifery service director]**

Theme three: supervision and feedback on data quality

Most in-depth interviews stated that there was supervision and feedback given, which helped to prepare an action plan for the gap given and to correct the gap mentioned as follows:

By the way, feedback is very good because they see things that you don't see. When they give you such a comment, if there is something good, it helps us to continue, and if there is a gap, it helps us to fix it, and it helps us to make an action plan to solve the gap based on that. **[KII 1, Dep coordinator]**

They will provide feedback as soon as they are supervised. If they come every three months, they will give feedback, if they come once in six months, they will also give feedback. For example, there is an element where they give feedback through nursing and midwifery and there is an element where they give it through the HMIS unit. When the performance monitoring team reviews the monthly report every time, and if the data is good, they give positive feedback to continue. I think having supervision and giving feedback is very beneficial for staff to improve data quality. **[KII 4, Hospital PMT member]**

Theme four: training and its impact on data quality

Many of the participants in-depth interviews mentioned that the training clear out different indicators which is new, this in turn enhance data quality and explained as follow:

I *think training is a must, like our hospital, because the indicators change every time, when the indicator changes, the registrations also change, because the data sources are registration points for many indicators. He or she doesn't know much about it, and the data he fills in is irrelevant. So indirectly it reduces our outcome and data quality.* **[KII 12, Quality officer]**

Training is very important. Now we have given disease classification, just as it was changed in July, we have given training to the data owner. Currently, the reporting format is better, that is, after we have given training. After we give HMIS training and before we give it, different indicators are very clear to them, it means after they take training. **[KII 2, HMIS officer]**

Training, when you take it, it will strengthen more because every time you take training, you will get something new and with that training, you will make a correct and valid report. Then we provide training to the staff through them, which means that it will be used to make data accurate at the end. [**KII 4, PMT member**]

6. Discussion

This study attempted to assess the level of data quality and explore the factors affect routine health information system data quality. The result indicated that the average data accuracy and consistency in this study were about 98% and 85% respectively, over all the level of data quality was about 91% with 95% CI [88.9 – 92.76]. In this study approximately 66 percent of health workers received training, around 58% of health professionals were supervised by higher officials and about 66% of health workers received feedback from higher official. In this research area Werabe comprehensive specialized hospital and Alem gebeya primary hospital were supported by CBMP/DUP Project of HIS Implementation.

Average data quality in terms of accuracy, completeness, timeliness and consistence was about 98%, 86%, 95% and 85%, respectively. Whereas the level of data quality at WCSH was around 90% and Tora primary hospital was about 89%, 90% for Kibet primary hospital and Alem gebeya primary hospital data quality was about 93% calculated by using average of four dimensions.

Study conducted at Addis Ababa city Administration revealed that overall data quality of health centers found to be 76%, at Hadiya zone, southern Ethiopia stated that overall data quality was 82%, and study done at Harari region, Ethiopia stated that, 51% departments have good data quality, these study results of data quality was lower than study conducted at Silte zone public hospital and average result of data quality in this study was in line with expected data quality at the national level, possible justification might be a Capacity Building and Mentorship Project (CBMP) implementation on most of study areas where regular technical and capacity-building support is provided in an attempt to strengthen HIS in the region and also information diffusion to the remaining hospitals as a result of these hospitals supported by project. Another possible justification might be, in this study public hospital only included, but most of study listed above included health posts and health centers, in addition to this, additionally dimension used, duration of study might be matter, and the variation of sample size might be possible reason for the defence (13, 45, 47, 57).

The accuracy of data in this health facility was 98%, which is in line with study conducted at Mozambique found that the indicators they used for their study had an accuracy range of 91-97% whereas the study conducted in Hadiya Zone (76%), data accuracy of study conducted in west Gojjam zone north west Ethiopia was 74%, study conducted in Harari region Ethiopia stated that over all data accuracy were 58% as well as Addis Ababa city administration (69%) and study conducted in Nigeria (76%), which is lower (21). The difference might be because of the difference in the type of facility, long duration of study and the difference in selected indicators to measure data accuracy.

Completeness is average percentage of report and source document, the overall Silte zone public hospital data completeness score were 86%, which is lower than in a study conducted at Addis Ababa the percentages for report (94%) and source document completion (96%) this might be Addis Ababa study used DHIS2 generated report completeness score, whereas in India (71%) and Harari region Ethiopia (60%) have complete data which is lower than this study, possible justification might be duration of study and change indicators used to measure completeness (13, 18, 45).

Another one is consistence which has been measured reportable data elements written on the register are exactly matched with what is recorded on individual medical records. In this study consistence of data elements between register and individual medical records were 85% whereas study done at Addis Ababa city, about 97% of reports from health centers were consistent which is higher than this study, the difference might be they were used aggregated data and way of measuring as well as duration of reported data used to assess consistence.(45)

In this study 95 percent of data were timely reported which was closed to the study done at Harari Region which was 94%, but higher than the timeliness reported from the other parts of Ethiopia-70% in East Wollega and in Addis Ababa city administration, the median report timeliness score was 33%, ranging from 0%-100%, this difference might be due to method of assessing timeliness that means these are assessed using the DHIS2 generated timelines report and the numbers of report reviewed (59).

In this study about 56% of health workers were received training regarding to HMIS activities, another study done at Hadiya zone Ethiopia found that about 52% of health workers get training

regarding to HMIS activities (47), receiving training of HMIS related activities have 1.6 time higher odd of having good data quality than participants who are not receive training which is important to create awareness and to have skilled human resources, other study conducted at eastern Ethiopia stated that having Trained staff have 2.3 times higher odd of having good data quality than these are not trained, (47). This is supported by the qualitative results as follow,

"I *think training is a must, like our hospital, because the indicators change every time, when the indicator changes, the registrations also change, because the data sources are registration points for many indicators. He or she doesn't know much about it, and the data he fills in is irrelevant. So indirectly it reduces our outcome and data quality*. **[KII 12, Quality officer]".**

Concerning supervision, regular supportive supervision with feedback is a key in addressing quality issues by helping to improve overall performance of HMIS especially for better achievement of data quality. (60)

In this study 67% of hospitals health professional had supervised by higher officials, whereas study conducted at Hadiya zone southern Ethiopia revealed that, more than half (63%) health centers supervised by their respective higher level in the last two quarters. (47) This difference might be due to difference in type of facility, whereas study done in Harari region Ethiopia stated, supervised health workers have a higher data quality than those who have not supervised and the studydone in Kenya was 79% were supervised, the finding of this study area was lower than that of a study done in Kenya and Harari region, the possible justification for the variation might be they incorporate different types of health facility including health post. (13, 61)

Getting supervision from higher officials have 1.7 times higher odd of having good data quality than these who have not get supervision from higher officials whereas in this study (66%) hospitals health professional get regular feedback by higher officials, feedback and supervision remains essential for achieving and maintaining improvements in data quality (45).

Providing regular feedback based on evidence to their staff through regular reports from supervisor had 1.7 times to odd of having good data quality than those who had not get regular feedback from supervisory. This is supported by the qualitative results in this study

"… *Feedback is very good because they see things that you don't see. When they give you such a comment, if there is something good, it helps us to continue, and if there is a gap, it helps us to fix it, and it helps us to make an action plan to solve the gap based on that.* [**KII 1, Dep coordinator]"**

Another participant emphasized by saying:

"*They will provide feedback as soon as they are supervised. If they come every three months, they will give feedback, if they come once in six months, they will also give feedback. For example, there is an element where they give feedback through nursing and midwifery and there is an element where they give it through the HMIS unit. When the performance monitoring team reviews the monthly report every time, and if the data is good, they give positive feedback to continue. I think having supervision and giving feedback is very beneficial for staff.* **[KII 4, PMT member]"**

From the findings of this study, 61% of study participants agreed that decisions and followup actions identified in PMT meetings based on presented data increase the quality of routine health information system but only 79% of service delivery point establish performance monitoring team, a study done in Addis Ababa reports that all sampled health centers had PMT. However, there were gaps in the consistency of the meeting and all sampled health centers have PMT (45), health care workers agreed on decisions are made and follow up actions identified in PMT meetings based on presented data have 1.7 times higher odd of having good data quality than others. This is supported by the qualitative results as follow,

"*There are performance monitoring team (PMT), so PMT report periodically, monthly data, weekly report, data we report at case team level, especially those monthly report data is evaluated at case team level by PMT, before being sent to our hospital or HMIS department. After they are confirmed, or after the quality of the data items are confirmed, they are sent to HMIS department this is one of the ways to keep data quality.*" **[KII 7, Quality officer]**

In this study, health workers agreed on user friendly or easily understandable registration and report format affecting data quality, had 1.9 times more likely to have good data quality than others, this finding is agreed with study done at west Gojjam zone, North West Ethiopia, those health workers who agreed to the complexity of the RHIS format affecting data quality had higher odd of good data quality than those who disagreed to the complexity of the RHIS format who are not affect data quality. This is supported by the qualitative results in this study. (60)

"… *I don't think there is a problem with the data register tool because the tool is easy to understand, if you need an explanation, you will find it written below in the registration or report format. Tool availability and easy understandability has a positive effect. Currently we have a tool that you can use to do all staff to reports and registers. And that has a positive effect on data quality.* **[KII 10, Nursing& midwifery service director]"**

Regarding motivation due to appreciation, appreciation may motivate health professionals toward RHIS activities are which in turn affecting the RHIS data quality, In this study about 60% of health workers were agreed on appreciation by supervisors or coworkers for work of data collection increase data quality, this is supported by literatures at study done Addis Ababa city administration, the result indicated that the motivation of service providers and health center data quality was found to be strongly positively correlated (45). The findings of this study showed that, these who are appreciated and valued by supervisors and co-worker have 1.6 times higher odd of having good data quality than others.

7. Conclusion and recommendation

Conclusion

The level of data quality at public hospital in Silte zone was 91% and from four dimensions of data quality, completeness and consistence were below 90% whereas data accuracy and timeliness is above 90%, The registration and report format easily understandability, receive training, get supervision, provide regular feedback, team's work appreciated and valued by supervisors and decisions are made and follow up actions identified in PMT meetings based on presented data were factors affect routine health information system data quality.

The data collection tools and its impact, data quality challenges and assurance mechanisms, supervision and feedback on data quality, and training and its impact on data quality were four themes identified during thematic an analysis.

Recommendation

For MOH and other supporting organization to fully transforming the paper based service registration to an electronic-based medical recording system in which it will reduceincompleteness and inconsistence of data and increase data quality.

For regional and zonal health bureau to increase supportive supervision and regular feedback to health professionals and to work on identified gaps.

For health facility level of managers to do on staff motivation and to bring sense of owner ship as well as uses of data at hospital and national level.

8. Strength and limitations of the study

Strength

Since this study were conducted by using both quantitative and qualitative data collection and tried to triangulate at discussion part of study.

This study were used four dimensions to state level of data quality and tried to include additional indicators to assess dimensions.

Limitation

The study was not able to include health posts and health centers to state overall zonal health facility data quality.

Since this study was a cross-sectional study, it was challenging to prove the temporal correlation in a cause-and-effect relationship.

8. References

1. Organization WH. Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies: World Health Organization; 2010.
2. 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. Ethiopian Journal of Health Development. 2021;35(1).
3. Kumar M, Gotz D, Nutley T, Smith JB. Research gaps in routine health information system design barriers to data quality and use in low-and middle-income countries: A literature review. The International journal of health planning and management. 2018;33(1):e1-e9.
4. Gimbel S, Mwanza M, Nisingizwe MP, Michel C, Hirschhorn L. Improving data quality across 3 sub-Saharan African countries using the consolidated framework for implementation research (CFIR): results from the African health initiative. BMC health services research. 2017;17(3):53-63.
5. Yarinbab TE, Assefa MK. Utilization of HMIS data and its determinants at health facilities in east Wollega zone, Oromia regional state, Ethiopia: a health facility based cross-sectional study. Med Health Sci. 2018;7(1):4-9.
6. AbouZahr C, Boerma T. Health information systems: the foundations of public health. Bulletin of the World Health Organization. 2005;83(8):578-83.
7. Drobac PC, Basinga P, Condo J, Farmer PE, Finnegan KE, Hamon JK, et al. Comprehensive and integrated district health systems strengthening: the Rwanda Population Health Implementation and Training (PHIT) Partnership. BMC Health Services Research. 2013;13(2):1-13.
8. Kumar M, Gotz D. System Design Barriers to HIS Data Use in Low and Middle-Income Countries: A Literature Review. Weekly Epidemiological Record. 2014;89(40):429.
9. Lippeveld T. Routine health facility and community information systems: creating an information use culture. Global Health: Science and Practice; 2017. p. 338-40.
10. Gimbel S, Mwanza M, Nisingizwe MP, Michel C, Hirschhorn L. Improving data quality across 3 sub-Saharan African countries using the consolidated framework for implementation research (CFIR): results from the African health initiative. BMC health services research. 2017;17:53-63.
11. Organization WH. Improving data quality: a guide for developing countries. 2003.
12. Bram JT, Warwick-Clark B, Obeysekare E, Mehta K. Utilization and monetization of healthcare data in developing countries. Big data. 2015;3(2):59-66.

13. Shama AT, Roba HS, Abaerei AA, Gebremeskel TG, Baraki N. Assessment of quality of routine health information system data and associated factors among departments in public health facilities of Harari region, Ethiopia. BMC medical informatics and decision making. 2021;21(1):1-12.
14. Roomaney RA, Pillay-van Wyk V, Awotiwon OF, Nicol E, Joubert JD, Bradshaw D, et al. Availability and quality of routine morbidity data: review of studies in South Africa. Journal of the American Medical Informatics Association : JAMIA. 2017;24(e1):e194-e206.
15. de Souza DK, Yirenkyi E, Otchere J, Biritwum NK, Ameme DK, Sackey S, et al. Assessing Lymphatic Filariasis Data Quality in Endemic Communities in Ghana, Using the Neglected Tropical Diseases Data Quality Assessment Tool for Preventive Chemotherapy. PLoS neglected tropical diseases. 2016;10(3):e0004590.
16. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health policy and planning. 2009;24(3):217-28.
17. Chiba Y, Oguttu MA, Nakayama T. Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. Midwifery. 2012;28(3):329-39.
18. Sharma A, Rana SK, Prinja S, Kumar R. Quality of Health Management Information System for Maternal & Child Health Care in Haryana State, India. PloS one. 2016;11(2):e0148449.
19. O'Hagan R, Marx MA, Finnegan KE, Naphini P, Ng'ambi K, Laija K, et al. National Assessment of Data Quality and Associated Systems-Level Factors in Malawi. Global health, science and practice. 2017;5(3):367-81.
20. Belay H, Lippeveld T. Inventory of PRISM framework and tools: application of PRISM tools and interventions for strengthening routine health information system performance. Chapel Hill. 2013.
21. Adejumo A. An assessment of data quality in routine health information systems in Oyo State, Nigeria. 2017.
22. Hotchkiss DR, Aqil A, Lippeveld T, Mukooyo E. Evaluation of the performance of routine information system management (PRISM) framework: evidence from Uganda. BMC health services research. 2010;10(1):1-17.

23. Glèlè Ahanhanzo Y, Ouedraogo LT, Kpozèhouen A, Coppieters Y, Makoutodé M, Wilmet-Dramaix M. Factors associated with data quality in the routine health information system of Benin. Archives of public health. 2014;72(1):1-8.
24. Belay H, Azim T, Kassahun H. Assessment of health management information system (HMIS) performance in SNNPR, Ethiopia. Measure Evaluation. 2013.
25. Glèlè Ahanhanzo Y, Ouendo E-M, Kpozèhouen A, Levêque A, Makoutodé M, Dramaix-Wilmet M. Data quality assessment in the routine health information system: an application of the Lot Quality Assurance Sampling in Benin. Health policy and planning. 2015;30(7):837-43.
26. Roomaney RA, Pillay-van Wyk V, Awotiwon OF, Nicol E, Joubert JD, Bradshaw D, et al. Availability and quality of routine morbidity data: review of studies in South Africa. Journal of the American Medical Informatics Association. 2017;24(e1):e194-e206.
27. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health policy and planning. 2009;24(3):217-28.
28. Organization WH. Everybody's business--strengthening health systems to improve health outcomes: WHO's framework for action. 2007.
29. Dagnew E, Woreta SA, Shiferaw AM. Routine health information utilization and associated factors among health care professionals working at public health institution in North Gondar, Northwest Ethiopia. BMC health services research. 2018;18(1):1-8.
30. <HMIS information use guide.pdf>.
31. <improving data quality.pdf>.
32. Simba DO, Mwangu MA. Factors influencing quality of health management information system (HMIS) data: The case of Kinondoni district in Dar es Salaam region, Tanzania. East African Journal of Public Health. 2006;3(1):28-31.
33. 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 services research. 2019;19(1):1-6.
34. Kirimi NS. Factors Influencing Performance of Routine Health Information System: the Case of Garissa Subcounty, Kenya: University of Nairobi; 2017.

35. Nwankwo B, Sambo MN. Can training of health care workers improve data management practice in health management information systems: a case study of primary health care facilities in Kaduna State, Nigeria. The Pan African Medical Journal. 2018;30.
36. Gebremedhin LT, Giorgis TW, Gerba H. Policies, delivery models, and lessons learned from integrating mental health and substance abuse services into primary health care in Ethiopia. FASEB BioAdvances. 2021;3(9):694.
37. Roadmap I. Ethiopian Federal Ministry of Health Information Revolution Roadmap. 2016. April.
38. Farhajuddin S, Khan S, Ramesh K, Khushk I, Saima H, Assad H. Assessment of completeness and timeliness of district health information system at first level care facilities in a rural district of Sindh, Pakistan. Pakistan Journal of Public Health. 2015;5(3):28-31.
39. Ndabarora E, Chipps JA, Uys L. Systematic review of health data quality management and best practices at community and district levels in LMIC. Information Development. 2014;30(2):103-20.
40. Solomon M, Addise M, Tassew B, Balcha B, Abebe A. Data quality assessment and associated factors in the health management information system among health centers of Southern Ethiopia. PloS one. 2021;16(10):e0255949.
41. Tayi GK, Ballou DP. Examining data quality. Communications of the ACM. 1998;41(2):54-7.
42. Chen H, Hailey D, Wang N, Yu P. A review of data quality assessment methods for public health information systems. International journal of environmental research and public health. 2014;11(5):5170-207.
43. Al Otaibi MN. Internet of Things (IoT) Saudi Arabia healthcare systems: state-of-the-art, future opportunities and open challenges. Journal of Health Informatics in Developing Countries. 2019;13(1).
44. Adane A, Adege TM, Ahmed MM, Anteneh HA, Ayalew ES, Berhanu D, et al. Exploring data quality and use of the routine health information system in Ethiopia: a mixed-methods study. BMJ open. 2021;11(12):e050356.
45. Haftu B. Assessment of routine health information system (RHIS) data quality and factors affecting it, Addis Ababa City Administration, Ethiopia, 2020: Addis Ababa University; 2020.
46. Gobena T, Shore H, Berhanie D, Kenay A, Wondirad Y, Ayanle M. Routine Health Information System Data Quality and Associated Factors in Selected Public Health Facilities of Jigjiga

Woreda, Somali Regional State's, Eastern Ethiopia. Ethiopian Journal of Health Development. 2022;36(1).
47. Getachew N, Erkalo B, Garedew MG. Data quality and associated factors in the health management information system at health centers in Shashogo district, Hadiya zone, southern Ethiopia, 2021. BMC medical informatics and decision making. 2022;22(1):1-9.
48. Abajebel S, Jira C, Beyene W. Utilization of health information system at district level in Jimma zone Oromia regional state, South West Ethiopia. Ethiopian journal of health sciences. 2011.
49. Ramamurthy KR, Sen A, Sinha AP. An empirical investigation of the key determinants of data warehouse adoption. Decision support systems. 2008;44(4):817-41.
50. Yusof MM, Kuljis J, Papazafeiropoulou A, Stergioulas LK. An evaluation framework for Health Information Systems: human, organization and technology-fit factors (HOT-fit). International journal of medical informatics. 2008;77(6):386-98.
51. Zodpey SP, Negandhi HN. Improving the quality and use of routine health data for decision-making. Indian journal of public health. 2016;60(1):1.
52. Deng Z, Lu Y, Wei KK, Zhang J. Understanding customer satisfaction and loyalty: An empirical study of mobile instant messages in China. International journal of information management. 2010;30(4):289-300.
53. Alqahtani A, Aljarullah AJ, Crowder R, Wills G. Barriers to the adoption of EHR systems in the Kingdom of Saudi Arabia: an exploratory study using a systematic literature review. Journal of Health Informatics in Developing Countries. 2017;11(2).
54. <Data_Quality_Review_1.pdf>.
55. Dagnew E, Woreta SA, Shiferaw AM. Routine health information utilization and associated factors among health care professionals working at public health institution in North Gondar, Northwest Ethiopia. BMC Health Serv Res. 2018;18(1):685.
56. Haftu B, Taye G, Ayele W, Biruk E, Habtamu T. A mixed-methods assessment of routine health information system (RHIS) data quality and factors affecting it. Addis Ababa City Administration, Ethiopia. 2020;2021.
57. Directorate E. Health data quality training module participant manual. Mathewos T, collaboration. 2018.

58. Mathewos T. Community health management information system Performance and factors associated with at health post of Gurage zone, SNNPR, Ethiopia [Internet]. University of Gondar and Addis Continental Institute of Public Health. 2015.
59. Kebede M, Adeba E, Chego M. Evaluation of quality and use of health management information system in primary health care units of east Wollega zone, Oromia regional state, Ethiopia. BMC medical informatics and decision making. 2020;20(1):1-12.
60. Shiferaw AM, Zegeye DT, Assefa S, Yenit MK. Routine health information system utilization and factors associated thereof among health workers at government health institutions in East Gojjam Zone, Northwest Ethiopia. BMC medical informatics and decision making. 2017;17(1):1-9.
61. Chiba Y, Oguttu MA, Nakayama T. Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. Midwifery. 2012;28(3):329-39.

Annex

Participant's information sheet and consent form

Dear respondent: My _____. I am here to interview you on behalf of Nigussie Dukamo he is a post graduate student of Hawassa University, College of Medicine and health science, school of Public Health. Currently, he is conducting a study for the purpose of partial fulfillment of Masters of Public Health in Epidemiology. I kindly request you to give me few minutes to explain you about the research. The objective of the study is to assess level of routine health information system data quality and identify associated factors at public hospitals in Silte Zone. In order to achieve the objective of the study, I will provide you a self –administered questioner which addresses socio demographic factors, Knowledge of RHIS, technical factors, organizational factors and behavioral factors. Your participation in this study does not involve any direct risk because it does not require taking any samples that pose risks and direct benefit on you, however; findings of the study will be useful in showing the current status of the routine health information system data quality and factors that affect it. This will in turn help to design intervention strategies that can improve the quality of routine health information system data for the better health of our community. Please be assured that all information collected will be stored strictly confidential and you do not required to mention your name on any part of the questionnaire. The interview may take only 20—30 minutes and if you are not willing to participate you can leave at any time. The result of the study will be presented here in the hospital as an aggregated data report, please be frank and choose your answer honestly.

Thank you for your time and cooperation.

Address of principal investigator: Name: - Nigussie Dukamo

Phone no. +251-939-15-15-66

Email: - nigusdukamo@gmail.com

Questionnaire

1. English version data collection questionnaires

Part I. Questions related to socio-demographic characteristics

	Name of facility_____Unit or department name_____	
1.	Position of person interviewed	1. Health care worker (staff) 2. Case team leader 3. HMIS officer 5. Others(specify_____)
2.	Age_____	
3	Years of employment _____	
4.	Salary per month_____ETB	
5.	Sex	1. Male 2.Female
6.	Educational status	1. Diploma 2. Bachelor Degree 3. Master's Degree 4.Others(specify_____)
7.	Professional category	1. Nurse 2. Midwife 3. Health Office 4. Laboratory 5. Health information& technology (HIT) 6. Pharmacist 7. Medical doctor 7. Other (specify)

We would like to know your opinion about how strongly you agree with certain activities carried out by you and your organization. There are no right or wrong answers, but only expression of your opinion on a scale. The scale is about assessing the intensity of your belief and ranges from strongly disagree (1) to strongly agree (5). You have to determine first whether you agree or disagree with the statement. Second decide about the intensity of agreement or disagreement. If you disagree with statement then use left side of the scale and determine how much disagreement that is: strongly disagree (1) or disagree (2) and circle the appropriate answer. If you are not sure of the intensity of belief or think that you neither disagree nor agree, then circle (3). If you agree with the statement, then use right side of the scale and determine how much agreement that is: agree (4) or strongly agree (5) and circle the appropriate answer.

Please be frank and choose your answer honestly. To what extent, do you agree or disagree with the following on a scale of 1-5?

Part II. Assessment of technical factors

S. N	Questions	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1.	There is standardized set of indicator.					
2.	The registrations and report formats are user friendly or easily understandable.					
3.	Do you agree that there is trained staff able to fill out format					
4.	The register or tally sheet are filled completely.					
Part III:- Knowledge of RHIS						
1.	HMIS Collects data from service and administrative records					
2.	HMIS Provides signals that can be reviewed frequently to monitor program implementation					
3.	HMIS is Used for decision making					
4.	HMIS is important for policy making and management decisions					
5.	HMIS is important for Monitoring and Evaluation of performance					
6.	HMIS data can be presented by using Charts, graphs and tables					
7.	HMIS is an integral part of Health Information System					
	Part IV. Organizational and Behavioral Assessment Tool					
1.	Have you received training in HMIS related activities in the last six months? 1. Yes 2. No					
	In the three months did you get supervision from higher officials? 1. Yes 2. No					
	If yes for Q 2, how many times supervised. 1. One time 2. Two times 3. Three times 4. More than three times					

5.	Staff can access computer 1.Yes 2.No					
6.	Staff can access internet 1.Yes 2.No					
S. N	Questions	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1.	Supervisors provide regular feedback to their staff through regular report based on evidence					
2.	Supervisors provide supportive supervision to the staff focusing on data quality					
3.	Supervisors emphasize data quality in monthly reports					
4.	Supervisors check data quality at the your unit					
5.	My work of collecting data is appreciated and valued by supervisors or co-workers.					
6.	Staffs perceive data collection as a useless Activity.					
7.	I feel that data collection or recording is not the responsibility of health care providers.					
8.	Staffs engage actively in all activities of the health information system.					
9.	Decisions are made and follow up actions identified in PMT (Performance monitoring team) meetings based on presented data					
10.	Staffs use HMIS data for day to day management of the facility					
11.	Staffs are empowered to make decisions based on HMIS data					
12.	Staffs display data for monitoring their set target by means of graphs and tables.					
13.	Your institution encourages to gather data to find the root cause(s) of the problem					

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
14.	Staffs use HMIS data to identify and predict or manage epidemics					
15.	Staffs use HMIS data for Drug supply and management					
16.	Staffs use HMIS data for community education and mobilization					
17.	I feel discouraged when the data that I collect or record are not used for taking action (either for monitoring or decision making)					
18.	Collecting information makes me feel bored					
19.	Collecting information is meaningful for me					
20.	I feel that the data I collect or I have are important for monitoring facility service performance.					
21.	Collecting information gives me the feeling that it is forced on me					
22.	Staffs manipulate the data for the sake of competition					
23.	I am negligent in keeping data quality					
24.	Staffs feel guilty for not accomplishing the set target or performance					
25.	Staffs are made accountable for poor performance					
26.	Staffs feel committed in improving health status of the target population					
27.	Staffs document their activities and keep records					

Name of data collector: _____Signature_____

Date of data collection: _____

Part V: - Check-List for units or department to assess data quality

1. General information			
	Name of facility_____Unit or department name_____		
1.	Is there designated staff (HMIS focal person)	1. Yes	2.No
2.	The responsibility for recording the delivery of services on source documents is clearly assigned (have written job descriptions)to the relevant staff	**1.** Yes	2.No
3.	Does the unit have electronic data base (computer software)? (If no, skip to 6)	1. Yes	2.No
4.	Is the electronic database (computer software) currently functional?	1. Yes	2.No
5.	Is there any system in place to prevent unauthorized changes to data?	1. Yes	2.No
6.	The service delivery point establish performance monitoring team	1.Yes	2.No
2. Data recording			
1.	Does this unit keep copies of the Routine HMIS reports, which are sent to the higher level?	1. Yes 2. No, if no, go to Q3	
2.	Count the number of Routine HMIS reports that have been kept at the unit for the last twelve months		
3.	Does this unit have HMIS registration book?	1. Yes 2. No	
1. Data Accuracy Check			
1. If yes to Q3, take the six indicators or data items from top ten disease and find the information in the registration book of the unit for the selected three months. If the unit does have not keep copies of the monthly report, obtain copies at the next higher level and complete the following.			

Months	November, 2022		December, 2022		January, 2023	
Disease or service reports	#from registration book	#from Report	#from registration book	#from Report	#from registration book	#from Report
Total score at each month = the ratio of register divided by report submitted to the next level						
Total score of accuracy is their average						

2. Data Completeness

4a: Source document completeness

Instruction

Data completeness on the registers for selected data element, Locate the register related with the selected indicators or units, then Count the total number of clients who are recorded on the register during the quarter (from November to January) and take 5% of the total recoded clients in the specific register by using systematic sampling method and agree on the most relevant and

a must to be filled data elements from the register, Locate the sampled MRN on the register and check for completeness of data for each of the sampled individuals.

A. Complete means that the source document contains all the data elements relevant to the selected indicator

B. Incomplete means that there are missing data elements relevant to the selected indicators.

Indicators			
Register			
Relevant data elements			
Sampled MRN	Complete	Incomplete	% percentage of data completeness on register = Total no of MRNs with complete data to total no sampled MRN
Total =	Total =	Total =	

4b: Report completeness

How many relevant data items does the department or unit need to report (expected) on the Routine HMIS report for the selected indicators, this number does not include data items for services not provided by this health department/unit if so indicating "zero".

Total number of data item reported from department or unit

Selected months	# Expected data items to report	# of reported data items	% percentage of data completeness on report
November			
December			
January			
Total			

Report completes = # of reported data items per #Expected data items

N.B. Completeness = Average of register and report completeness.

4. Timeliness

1	Does this department/unit submit a report on time?	1. Yes	0. No

If yes to Q8, check the date of submission for three months monthly reports to answer Q9 (November, December 2022 and January 2023), If the unit does have not keep copies of the monthly report, obtain copies at the next higher level.

2	Months	Within the reporting period interval
	November	1. Yes observed 2. No
	December	1. Yes observed 2. No
	January	1. Yes observed 2. No

N.B:- Timeliness of report =total number of service or disease reports sent timely / Total number of expected reports

Timeliness of report =

5. Consistence

Consider the already agreed data elements (at Step 1 of the register completeness) and locate the individual medical records of sampled MRN from the MRU, check if the reportable data elements written on the register are exactly matched or not matched with what is recorded on individual medical records.

"Matched" is defined as when all of the selected data elements that are recorded on the registers for the sampled individual are also recorded the same on the individual medical records.

"Not Matched" is defined as when at least one or more of the selected reportable data elements that are recorded on the registers for the sampled individual is not exactly the same as what is recorded on the individual medical records. In addition, if the individual medical record is not physically available, then it is also considered as "Not matched".

Sampled MRN	Data elements from individual medical records		
	Matched	Unmatched	Data Consistency level individual medical register against register
			Data consistence = Total MRNs with matched / Total Sampled MRN

Part VI: KII questions for qualitative study adapted from PRISM tool and reviewing different literatures.

Good Morning/Afternoon dear respondent:

My name is_____. I am here on behalf of Nigussie Dukamo who is a post graduate student of Hawassa University, College of Medicine and health science, school of Public Health. I kindly request you to give me few minutes to explain about the study. The title of the study is the level of routine health information system data quality and associated factors at public hospital in Silte Zone, Central region Ethiopia.

Participation in this study does not involve any direct risk because it does not require taking any samples that pose risks or benefit on you. However; findings of the study will be useful for your facility in showing the current status of RHIS data quality and factors that affects it. This will in turn help to design intervention strategies that can improve routine health information system data quality by concerned body. Please be assured that all information collected will be stored strictly and you do not required to mention your name on any part of the questionnaire. You are selected purposively considering your knowledge regarding the topic of study in your facility better than others. The interview may take around 30-40 minutes and if you are not willing to participate you can leave from the study at any time. Findings and recommendations of the study will be given to your hospital at the end of the study. Please be frank and discuss your answer honestly.

Address of principal investigator: Name: - Nigussie Dukamo

Phone No. +252-9-39-15-15-66,

Email: - nigusdukamo@gmail.com

Are you willing to participate in this study? 1. Yes 2. No

If yes, proceed to interview guided questionnaire below; if no, thank you!

Key informant interview questionnaire

General information

Facility name_____Position or role _____

Educational background_____

1. Can you please describe overall routine process of collecting data in your health facility and what is your understanding of data quality?

2. Are there any challenges and facilitators you encounter in data quality and do you work with anyone to ensure data quality in your health facility? Please describe.

3. Do you use any data quality assurance mechanisms and how do you ensure data quality in your health facility? Please describe.

4. Would you please tell me that how training facilitates or affects RHIS data quality in your facility and how often your staffs have received training on RHIS?

5. Would you please tell me that your facility is supervised by concerned body regarding RHIS data quality? How often?

6. Do you think receiving periodic and regular feedback from supervision could facilitate or hinder RHIS data quality in you facility? How?

7. Would you please tell me how the negligence or attitude can facilitate or hinders the RHIS data quality?

8. Would you tell me please, how the availability of HMIS tools and Friendliness of reporting format enhances or affects the RHIS data quality in your facility?

9. Dear respondent, I finished my questions but if you have something want to add regarding RHIS data quality; you can. Otherwise thank you so much.