FACTORS INFLUENCING DATA QUALITY IN ROUTINE HEALTH INFORMATION SYSTEM AT HEALTH FACILITIES IN MARIDI COUNTY-SOUTH SUDAN

LUBANG DENSON

SHS/MPH/4093-2/2021

A RESEARCH THESIS SUBMITTED IN PARTIAL FULFILMENT FOR THE DEGREE OF MASTERS IN PUBLIC HEALTH (EPIDEMIOLOGY), DEPARTMENT OF COMMUNITY HEALTH, SCHOOL OF PUBLIC HEALTH

AMREF INTERNATIONAL UNIVERSITY

JUNE 2024

DECLARATION AND APPROVAL

Declaration by Candidate:

This thesis is my original work and has not been presented for a degree in any other university or any other award.

Signature:

Lubang Denson

SHS/MPH/4093-2/2021

3rd June 2024

Declaration by Supervisors:

This thesis has been submitted with our approval as university supervisors.

Signature:

Sr. Dr. Margaret Wandera Nyongesa

Technical University of Kenya

3rd June 2024

Signature:

Dr. Tobijo Denis Sokiri Moses (MPH, Ph. D)

Technical Adviser, The Rescue Initiative-South Sudan (TRI-SS)

3rd June 2024

ACKNOWLEDGEMENT

Firstly, I’m grateful that God has allowed me to pursue this public health master's degree. I am writing to express my gratitude and acknowledge the efforts and role played by my supervisors, Sr. Dr. Margaret Nyongesa and Dr. Tobijo Denis Sokiri Moses, in guiding and supporting the study and not forgetting Dr. Michel Mutabazi, who supported me during the corrections. Your efforts in this study are invaluable.

In a very special way, I am grateful to Amref Health Africa for covering part of my tuition while studying at AMIU.

My great family members, Biizu Stella, and Justin Morris Dumba, continued to provide support tirelessly to ensure I completed this study; I am grateful for everything, not forgetting everyone who helped with this study in any form.

ABSTRACT

Background: The study examined the factors affecting data quality in Maridi County, South Sudan, aiming to improve resource forecasting and equitable health service delivery. The lack of data has led to drug shortages and late reporting of morbidity data, causing the Ministry of Health to use a push system for resource allocation. This system is problematic, as many health facilities struggle to meet set targets.

Methods: A descriptive cross-sectional study was conducted on 12 functional healthcare facilities in Maridi County, with 106 respondents selected using simple random sampling. The researcher used SPSS version 25 for descriptive analysis, factor analysis to understand the relationship between independent and dependent variables, and thematic analysis to generate critical perspectives on data quality, focusing on behavioral, organizational, and technical aspects.

Results: The study found that data quality is impacted by insufficient motivation, negative staff attitudes, excessive workloads, lack of cooperation, personnel insufficiency, inadequate supervision, feedback, and training, with over 50% of variables showing weak to strong correlations in the factor analysis results. The regular feedback from the CHD ($r=0.683$, $p=0.007<0.05$), review meetings on data quality ($r=0.522$, $p=0.041<0.05$) & years of work experiences ($r=-0.555$, $p=0.031<0.05$) found to have a significantly strong correlation with data quality. In contrast, the other variables have an insignificant correlation with data quality.

Conclusion: The study reveals challenges in data quality, such as lack of motivation, work overload due to inadequate human resources, poor supervision of health facilities, feedback, insufficient training, and lack of reporting tools. It suggests that several strategies can be used to achieve high-quality data, including staff motivation, hiring more health workers to fill human resource gaps, frequent facility supervision, feedback provision, staff training on HMIS, and provision of data collection and reporting tools by the County Health Department.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AFSS: Andria Foods South Sudan

AMIU: Amref International University

ANC: Antenatal care

CHD: County Health Department

CUAMM: Doctors with Africa

DHIS2: District Health Information System version two

EWARS: Early warning alert and response

HIS: Health Information System

HMIS: Health Management Information System

KII: Key informant interview

M&E: Monitoring and Evaluation

MOH: Ministry of Health

MPH: Master's in Public Health

OPD: Outpatient Department

PHCCS: Primary Health Care Centers

PHCUS: Primary Health Care Units

PRISM: Performance of routine information system management

RHMIS: Routine Health Management Information System

DEFINITION OF KEY TERMS

Behavioral factors Behavioral factors are elements such as employee competence, skills for analyzing the data's quality, solving issues related to tasks involving HIS, competence in HMIS activities motivation, and the attitude of staff toward health information systems (Kleiman et al., 2020 & Chanyalew et al., 2021)

Completeness Data are accurate when the information recorded reflects the truth (CDC, 2020)

Data quality Suitability of data for use in terms of timeliness, completeness, and accuracy (Shama et al, 2021).

Data of good quality Data that are fit to be used or satisfy requirements for use (Haug et al, 2011)

Organizational factors Information about the organizational culture, available resources, and the major contributors' positions and duties at each healthcare system level.

Timeliness Timeliness of data is when all anticipated reports are submitted before a predefined deadline (WHO, 2017)

Technical factors Procedures, systems, and tools used for collecting data that impact health care data.

CHAPTER 1: INTRODUCTION

1.1 Overview

This chapter provided a detailed background to the problem, the problem statement, research questions, the broad objective, and specific objectives, justification for the study and its significance, the scope, limitations, and underlying assumptions.

1.2 Background Information for the Study

This study explored the interrelated factors that influence data quality in Maridi County, Western Equatoria State, South Sudan. It is conceived on the background that, as a young country, struggling with multiple health system challenges, effective management of health data will improve resource forecasting and equitable health service delivery. The question of data remains a sticking point for donors, the Ministry of Health and implementing partners. In most cases, critical shortages of drugs in many health facilities have been blamed on the absence of consumption reports hence the national Ministry of Health of South Sudan was compelled to employ the push method for drug distribution due to inaccurate and delayed reports of morbidity data. Unfortunately, such a system is so problematic that many health facilities struggle to meet set targets.

An effective health system must have the ability to provide health information. Global pledges to enhance health outcomes and systems have resulted in enhanced health management information systems (HMIS) which are used in program planning and decision-making at all levels of the health system, HMIS generates data on the availability of health services and the general health of the population. All other aspects of the health system's decision-making processes should be guided by timely and high-quality data from an information system (Li et al., 2018).

The challenges affecting data quality in routine health information systems cut across globally in Lumbini province, Nepal data quality was assessed through completeness and timelines and the results revealed overall completeness was found within 98% to 100% while timeliness ranged from 94% to 96% (Sanjel et al., 2024). Similarly, in Myanmar, 30.4% of Routine health information system data is of good quality, with data completeness of 30.4% and reporting timeliness of 31.9% (Hlaing et al., 2022). In Oyo State, Nigeria, data completeness was 77.3%, and data timeliness was 14% (Adejumo, 2017). The results further showed that workers, infrastructure, and data collection and management procedures are the primary elements affecting data quality meanwhile in the neighboring nation of Kenya, an evaluation of HIV data reporting performance reveals that in 2017, Timeliness was 83% and completeness was 97%, however in 2018, there was a substantial reduction in timeliness by 11% and completeness by 13% (Ngugi et al., 2020).

Unpublished data from the DHIS2 in South Sudan for 2021 revealed that completeness was 52.1% and timeliness 46.5%. In the Western Equatoria State of South Sudan, the completeness of data was 52.9%, while timeliness was 51.6%. In Maridi County, completeness of data was 76.1%, with timeliness at 72.8% (SSD DHIS2, 2021). The South Sudan Data Quality performance targets are 90% for completeness and 85% for timeliness (Mathewos, 2015).

An assessment conducted in Maridi County from September 12 to 16, 2022, revealed discrepancies in the reported data for some selected data elements, such as penta3, outpatient consultation, ANC first and fourth visits, and skilled deliveries. The analysis revealed that there was a prevalence of over-reporting or under-reporting in all health facilities. Due to these patterns, the accuracy, completeness, and timeliness of data in Maridi County have been compromised, leading to a performance that falls short of the national targets.

Although structurally well-developed, the implementation of HMIS in South Sudan remains weak. Key limitations to poor quality data at the facility level are driven by multiple factors such as excessive and complex reporting systems, a lack of digital technology, low motivation among healthcare workers, a lack of feedback, low pay, unfavorable working conditions, a lack of training, and a lack of data management skills, all contribute to poor quality data at the facility level (Shamba et al., 2021). Poor health data quality results may misdirect decision-making regarding allocating resources, and reliable regular healthcare data are necessary for the whole health information system to succeed (Kuyo, 2018).

1.3 Statement of the Problem

Maridi County is among the counties benefiting from the Health Pooled Fund. A multi-donor funding mechanism that supports Primary and Secondary health care in seven of the ten states of South Sudan. The county has one of the leading health training schools in South Sudan. The Health Pooled Fund, through its partners, supports health systems strengthening, including the HMIS

Before the 2016 subnational unrest, Maridi County had 23 operational health facilities. Currently, only 12 health facilities can be characterized as fully functional. The DHIS2, an upgraded version of the program, has transitioned to a cloud-based platform. Owing to the absence of internet connectivity in these facilities, the County Health Department utilizes a combination of manual and electronic HMIS. Data entry occurs at the County Health Department, where development partners are co-located and have internet access.

It is important to remember that planning and monitoring the effectiveness of health systems depend heavily on routine data collection, analysis, interpretation, and input from healthcare facilities. Moreover, if routine healthcare data meets a certain quality standard, it might be utilized for alternative objectives. Morbidity statistics can be used to estimate the burden of diseases and help shape healthcare policies (Roomaney et al., 2017). However, in underdeveloped countries, daily usage of data for decision-making still needs to be improved, primarily because of insufficient data Wandera et al. (2019). Hence, quality data are important because they help health managers/decision makers to make appropriate decisions regarding allocating resources meaning lack of data quality affects resource allocation.

In 2019, the South Sudan Ministry of Health and its implementing partners created easily understandable tools for collecting regular health information. Under the guidance of the National Ministry of Health, a roll-out training was organized for health facility staff, concentrating especially on those in charge of managing the data. Additionally, refresher training was provided to individuals who had previously received training.

Notwithstanding the many efforts, certain health facilities consistently provide monthly reports that are both inaccurate and incomplete and occasionally delayed (SSD DHIS2, 2021). Thus, the study aimed to identify the factors affecting data quality in health facilities in Maridi County of South Sudan.

1.4 Research Questions

1. What are the technical factors affecting the data quality in routine health information systems at all health facilities in Maridi County?
2. What organizational factors influence data quality in routine health information systems at all health facilities in Maridi County?
3. What behavioral factors influence data quality in routine health information systems at all health facilities in Maridi County?

1.5 Research Objectives

1.5.1 General Objective

The purpose of this study was to identify factors associated with data quality in routine health information systems in Maridi County, Western Equatoria State, South Sudan.

1.5.2 Specific Objectives

1. To determine the technical factors that affect data quality at all health facilities in Maridi County.
2. To examine the organizational factors that influence data quality at all health facilities in Maridi County

3. To assess the behavioral factors associated with data quality at all health facilities in Maridi County.

1.6 Justification of the Study

Very few studies examined the quality of routine health data in South Sudan. There are equally limited studies that evaluated the effectiveness of the routine HMIS in this country. South Sudan has only one open-access journal, which depends on international volunteers for article review. Most of the published data is individual-based, with limited institutional-based publications. It is, therefore, prudent to say that the publication space in this country is very limited, and so is data.

However, health programming now relies more on research findings as a basis for evidence-based policy and decision-making. Health institutions are increasingly interested in utilizing this data to enhance program design. An exemplary instance was the study's findings on the prevalence of epilepsy linked to onchocerciasis in Mundri County. Additionally, regular health information systems must be of high quality to guarantee proper healthcare delivery and the creation of suitable health policies (Glele et al., 2015).

The quality of regular health information data gathered from health facilities depends on the health system's ability to operate effectively and policymakers' ability to assess the results of systemic initiatives to enhance population health (Lemma et al., 2020).

The study sought to draw associations between the prevalence of epilepsy among populations that received annual distribution of ivermectin, a drug used as chemoprophylaxis for the prevention of onchocerciasis. The government of South Sudan had set such an approach to reduce the disease burden. However, morbidity data continued to report new cases. Therefore, the study recommended a bi-annual distribution to strengthen its effect in lowering the prevalence of epilepsy (Jada et al., 2022). The recommendation for implementation was picked positively by the National Ministry of Health and has since shown good results in reducing the disease burden. This example is a testament to using research findings to influence policy.

Further limitations to information arise from data protection policies driven by donors, such as restricted access to the DHIS2 system and supervision reports from the state ministry and donors. The presence of public and private healthcare institutions in Maridi County means that research and publications are missed opportunities. Motivated by this gap, the researcher chose to undertake this study to shape leadership in research and to add more information locally and to the prevailing global literature.

1.7 Significance of the Study

The need for accurate health data to inform policy and program design is not an indulgence but rather a moral decision. In general, there is a consensus worldwide that routine healthcare data have the potential to help the adoption of evidence-based interventions. The aforementioned data are extensively utilized in patent identification, result evaluation, and ensuring fairness in providing health services (Xie et al., 2023).

South Sudan employs a hierarchical system of health data administration, where counties are responsible to state data managers, who are answerable to their national counterparts. The extended data pathway hinders the prompt provision of feedback and allocation of resources. It also loosens the health ministry's oversight role to improve the quality of the data.

States may increase their involvement in planning due to a better understanding of their challenges. However, to achieve effectiveness, it is essential to have superior-quality data, which is obtained through the implementation of standardized procedures that are glued to the technical aspects of the system, improved organization, and appropriate behavior of the data system. Most research concerning data quality in South Sudan tends to be generalized, typically focusing on the national perspective, even though the health system is decentralized.

The study provided a comprehensive understanding of the variables that impact data quality in healthcare facilities in Maridi County. The state and County Health Departments can utilize this study to enhance data quality, leading to a more effective allocation of resources. By aligning with standard data quality practices, healthcare organizations can minimize resource wastage, increase staff motivation, and improve advocacy. Future research on a comparable topic can use this study as a benchmark.

1.8 Scope of the Study

This study investigated the technical, organizational, and behavioral aspects that contribute to the potential for enhancing data quality in Maridi County, South Sudan. The research was restricted to 12 health facilities based on pragmatic considerations regarding time, resources, security, and functionality. This all-inclusive study had equal considerations for women, men, and people with disability who worked at the selected facilities. The methodological scope was cross-sectional, applying quantitative and qualitative study techniques. The tools selected were a structured questionnaire for the quantitative aspect and Key Informant interviews for the qualitative inquiry. The scope of data analysis included univariate analysis, bivariate analysis for quantifiable information and qualitative data, and thematic content analysis. Quantitative data were collected before the qualitative data as a strategy to triangulate the findings.

1.9 Assumptions of the Study

1. The initial assumption was that the security situation would remain normal in Maridi County, allowing access to all the health facilities where the study would occur. Indeed, there were no reported security incidents that affected data collection.
2. The second assumption was that the selected health facilities would remain open and operational during data collection and that the health workers selected for the interview would be willing to participate. This study demonstrated an excellent response rate of 100%, as projected.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This part summarizes pertinent findings from similar studies on factors influencing data quality. It identifies gaps and compares results in previous studies. The literature review was guided by a literature search strategy, and keywords, data quality, regular health data, and HMIS were used to identify themes that were relevant to the study. Google Scholar and Semantic Scholar, among other search engines, were used to gather information.

2.2 Theoretical Framework

2.2.1 The Systems Theory

This study is grounded on the systems theory (Fig. 1), which was developed by Ludwig Von Bertalanffy, an Australian biologist, in the 1940s. According to Hooker (2011), this theory posits a framework for examining any collection of components working together to achieve a goal. Systems designers of the HMIS seek to understand the internal and external factors of data quality and the feedback mechanisms involved in communicating results. These feedback loops occur when a system eventually feeds back into itself in a circular fashion because its outputs influence its inputs (Social Work Theories, 2023). This theory aligns effectively with the goals of this study, which aims to investigate the related elements that impact the regular health information system. The study comprises three distinct components: technological, organizational, and behavioral aspects, the interplay of which determines the quality of medical data.

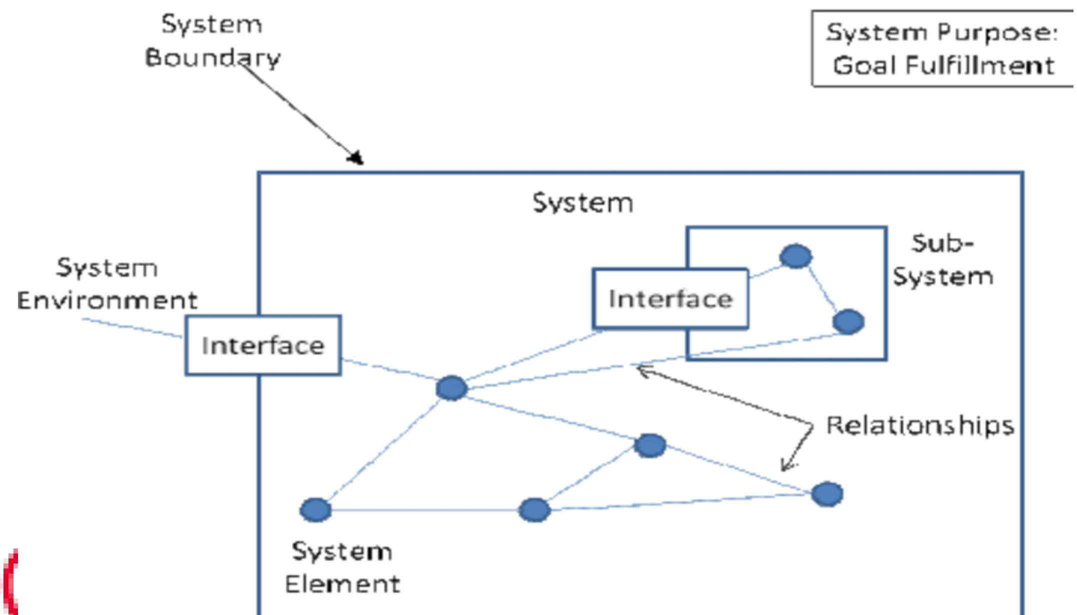

Figure 1: The Systems Theory (Matok & Brown, 2008)

2.3 Review of Related and Empirical Literature

2.3.1 Definition and Evolution of the Health Management Information Systems (HMIS)

The practice of collecting, storing, and utilizing health data is not new. In the 1960s, medical and Hospital Information Systems were introduced to facilitate administrative and medical responsibilities. These technologies had very limited scope to optimize financial returns and streamline the process of admitting patients.

The modern-day HMIS evolved from HIS. A concept popularized by Lippeveld as a comprehensive endeavor to gather, analyze, present, and utilize knowledge and data on health to impact policy-making, program implementation, and research (Epizitone et al., 2023).

The World Health Organization refers to the HMIS as the production of information to enable healthcare system decision-makers to identify obstacles and needs, decide on health policies, and allocate limited resources efficiently (WHO, 2008).

2.3.2 Performance of the Routine Health Management Information Systems

A proficient HMIS acquires precise, uniform, and pertinent data promptly to facilitate enhanced planning and monitoring of health activities (Meghani et al., 2022). However, despite clear universal guidelines, many countries have performed below country targets. The World Health Organization's (WHO) global assessment of the HMIS shows that approximately 40% of countries exhibit problematic practices in data quality assurance, and a significant number of countries lack the technological capacity to verify the accuracy of health data. The report also acknowledges that many lower- and middle-class countries depend on outside assistance for technical support and the infrastructure needed to build a strong HMIS.

Relatively low performance (37% and 29%, respectively) was found in evaluating the HMIS in the Indian state of Kerela. The accuracy, comprehensiveness, and punctuality of the procedures in the facilities were 79%, 79%, and 88% respectively. The level of proficiency in data analysis was 35%. The general degree of assurance in HMIS-related responsibilities was 69.4%, while the level of proficiency was 58%. According to Harikumar (2012), the percentages regarding the management duties of planning, monitoring, training, governance, and quality control at the facility level were 13.2%, 43.4%, 5.3%, 28.4%, and 44.7%, respectively.

The Health Information Management System in most African nations exhibits a significant performance deficiency across various measures. Consequently, the data quality in these countries is persistently inadequate to the extent that Musa et al. (2023) described it as patchwork. This is due to insufficient data availability and frequently poor quality.

The data completeness percentage in Oyo State, Nigeria, was significantly higher at 77.3%. However, this did not align with its accuracy rate, which was a mere 14% (Adejumo, 2017). The prevalence of such inconsistencies indicates the presence of systemic problems in their regular health information management system.

In Sudan, the HMIS's performance has increased but then stagnated. The reporting rates through the DHIS2 system increased from 30% in 2016 to 64% in 2020 but have since remained stable at 61.5% recorded in 2018 (WHO, 2022). Similar to South Sudan, Sudan has plunged into a civil war that is reversing the advancements made in bolstering its healthcare system.

Rumisha et al. (2020) found that tally sheets were used in just 77.8% of basic health facilities in Tanzania. The instruments in the dispensary, health center, and hospital had availability rates of 91.1%, 82.2%, and 77.8%, respectively. Nevertheless, the metropolitan districts demonstrated a very low tool availability rate of 65%. Occurrences of inaccurately filled out paperwork and insufficient adherence to coding guidelines were observed.

According to Teklegiorgis et al. (2016), the overall data quality in Eastern Ethiopia's departments and/or units was 75.3. The data quality was assessed to be lower compared to the national standard. Health units demonstrated low-quality data compared to hospitals and health centers. Ethiopia is a vast country with decentralized governance structures, which means resources and efforts are not equally distributed.

According to Rumunu et al. (2022), the nationwide implementation of electronic reporting for Early Warning Alert and Response (EWARS) in South Sudan complemented the DHIS2. Compared to a baseline of 54% on both timeliness and completeness of reporting in 2019, the weekly reporting improved to 78% and 90% by week 39 of 2020. Unfortunately, most of these achievements are driven by donor funding, and the lack of government commitment to sustain these improvements means these efforts are not sustainable.

2.3.3 Factors Associated with the Performance of Routine Health Information System

Looking back at the structure of the RHMIS, three main domains are derived and are popularly used to characterize the factors associated with its performance. These factors, documented in several studies, include technical, behavioral, and organizational factors (Sako et al., 2022). However, if these influencers are unpacked, they yield determinants such as data collection tools, standard indicators, and trained data team; feedback and supervision; motivation, level of knowledge, and attitudes of staff (Nguefack-Tsague et al., 2020)

2.3.3.1 Technical Factors.

The technical factors influencing data quality include systems, forms, procedures, techniques for gathering data, data collection tools, standard indicators, and trained staff (Kirimi, 2017; Dagnew et al., 2018). According to Wude et al. (2020), data quality is strongly influenced by the availability of trained staff and a standard set of indicators. This was a qualitative study, and it is challenging due to the absence of measurable evidence to gauge the extent to which these factors influence data quality.

A study in Myanmar discovered that multiple reporting, inexperienced personnel, and a deficiency of reporting tools are among the technical issues influencing data quality (Hlaing et al., 2022).

Another qualitative study in Uganda involving sixteen interviews with key informants and a workshop with several stakeholders established a link between the quality of the data and the availability and complexity of reporting tools (Wandera et al., 2019). However, due to the exclusive utilization of qualitative approaches, it was impossible to assess statistical significance.

Similar technical gaps were found in Kenya; the insufficient competence of staff, the presence of multiple Health Information System tools, and the lack of computers affected data quality in Tharaka Nithi County, Kenya (Mucee et al., 2016)

2.3.3.2 Organizational Factors.

According to Glette et al. (2021), training personnel involved in health data management, feedback on data quality, supportive supervision, and working conditions for health personnel are examples of organizational factors. Lemma et al. (2020) suggest that capacity-building measures, such as training, data quality assessment, and feedback provision to healthcare facilities, aid in raising the standard of the data. Their research attempts to offer a more comprehensive grasp of the utilization and accuracy of routine health data in middle-class and underdeveloped nations.

A study by Moloko et al. (2022) in Tshwane, South Africa, revealed that training, supportive supervision, and enough human resources influence the quality of data. These findings corroborated the research conducted in Northwest Ethiopia by Afework (2022), which indicated deficient feedback systems, insufficient human resources, and inadequate training as barriers to data quality.

In a cross-sectional study conducted in Kenya, Cheburet et al. (2016) found that support supervision positively impacted data quality. They recommended addressing these organizational aspects to ameliorate the data's quality.

The results of a related study by Shiferaw et al. (2017) in Gojjamzone, Northeast Ethiopia, showed that supportive supervision, HMIS training, and providing feedback to the health facilities were significantly associated with data quality. These results align with those of Tulu et al. (2021) in Ethiopia, which found that supportive supervision and HMIS training were significantly associated with data quality.

2.3.3.3 Behavioral Factors.

Behavioral factors are elements such as employee competence, skills for assessing the data's quality, solving issues related to tasks involving HIS, competence in HIS activities motivation, and the attitude of staff toward health information systems (Kleiman et al., 2020; Chanyalew et al., 2021)

Glèlè Ahanhanzo et al. (2014) identified worker demotivation and low capability as factors contributing to poor data quality in everyday operations related to health information systems. Hlaing et al. (2022) suggested that work burdens affect healthcare data quality because human resource shortages can result in work overload. This study supports their theory that the competency of healthcare workers, as measured by their education and involvement at work, is related to the quality of the data.

According to Moses et al. (2019), the efficiency of data collection, processing, and interpretation in South Sudan is impacted by the shortage of competent personnel at healthcare facilities. Inadequately trained people may fail to gather specific or erroneous data, compromising the overall quality of routine health information. Conversely, Haftu et al. (2021) identified the absence of a skilled HMIS focal person and a lack of motivation for HMIS responsibilities as obstacles to ensuring data accuracy in Ethiopia.

2.3.4 Challenges in Implementing an Effective Routine Health Management Information System

Given that HMIS is an integral part of the health system, its problems are consistent with its larger issues. These hurdles encompass inadequate healthcare personnel, limited healthcare funding, governance impacted by political instability, and numerous other obstacles. A study by Maiga et al. (2019) uncovered recurring problems with the accuracy and reliability of the data associated with target population estimates and served as the foundation for calculating coverage figures in 14 countries.

Related studies in Kenya and Ethiopia found inadequate staffing, the design of tools for gathering data, as well as the absence of essential resources as the main challenges in

Kenya, while lack of register books, the intricacy of the indicators, and the choice of denominators depending on population estimates as key challenges in Ethiopia (Adane et al., 2021). Comparatively, Tilahun et al. (2022), in a study in Ethiopia, recorded challenges such as HMIS staff capacity, HMIS code, excessive data sources, inadequate data quality assurance, charting, and data transfer guidelines.

2.3.5 Performance of Routine Information System Management (PRISM) Framework

MEASURE Evaluation, a technical body with substantial experience in health management information systems research, created the PRISM framework. The tool underwent a pilot phase in 2011 and was subsequently upgraded in 2018, incorporating additional features that have contributed to its widespread use in the present day. The process of refining it required several years of meticulous investigation. Kawakyu (2023) investigated whether modifications to the Routine Health Information System occurred before or after the intervention. The study concluded that data were crucial for enhancing the application of RHIS. The framework is equipped with instruments for collecting data, instructions for methodology, and methods for analyzing data. The regular health information framework's performance is affected by internal and external factors/inputs and outputs, as postulated by the results-based monitoring and evaluation paradigm. The components were categorized into technical, organizational, and behavioral factors, forming this study's structural foundation (Aqil et al., 2009).

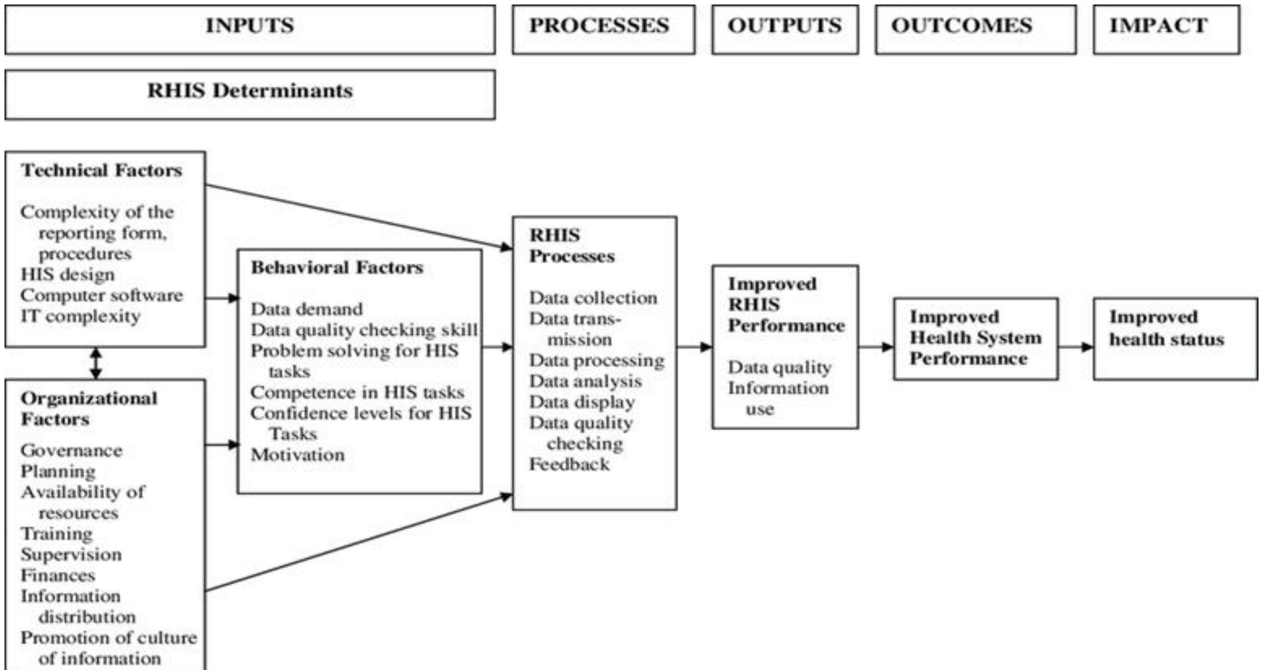

Figure 2: The PRISM Framework (Aqil et al., 2009)

The enhanced PRISM model was evaluated in multiple countries to determine its effectiveness in enhancing the performance of the regular health information framework. One of the main discoveries in Uganda was that the culture of information, proficiency in RHIS tasks, and motivation all contributed to improved performance. Similarly, the presence of self-assurance and the accessibility of personnel were linked to a rise in data usage. (Hotchikiss, 2010)

2.4 Identification of Knowledge Gap

The knowledge gap was determined by thoroughly examining existing literature and desk research. In terms of literature, the search for routine health information in South Sudan yielded significant outcomes on Google and Semantic Scholar search engines. However, the majority of these studies were deemed irrelevant or lacked the necessary specificity to be included. Nevertheless, a thorough examination conducted in collaboration with development partners and the County Health Department determined the specific areas of inquiry for acquiring fresh insights. Only one study in Maridi County was relevant to the research issue and specifically focused on data on family planning contained in the HMIS.

2.5 Conceptual Framework

A conceptual framework is utilized in research to delineate potential alternatives for illustrating the favored methodology (Suman, 2014). They serve as the framework for constructing the research questions and analysis. The researcher explored and established the definitions of the topics and elucidated the connections between them. The study categorized the variables affecting data quality in RHMIS into two sets of variables: independent variables and dependent variables with an intervening variable (Fig. 3)

2.5.1 The Independent Variables

An independent variable in a study may be changed to investigate its effects. It stands alone from any other variables (Bhandari, 2023). Hence, the fundamental variables constituting this investigation's essence were technological, organizational, and behavioral elements. The variables were derived from the PRISM paradigm.

2.5.2 The Dependent Variable

The variable under investigation being measured or evaluated is known as the dependent variable and is influenced by modifying the independent variables, as stated by Cherry (2022). The study assessed data quality as the dependent variable, evaluated based on the timeliness and completeness of the data obtained from health facilities and reported to the County health department. Data quality management is an essential part of the data management process. It involves efforts to enhance data quality, commonly connected to data governance initiatives, which aim to preserve uniform data layout and use within an organization (Stedman & Vaughan, 2022).

2.5.3 Intervening Variable

An intervening variable is a scientific concept that describes relationships between independent and dependent variables, adjusting for changes in the dependent variable due to the independent variable (Shaw, 2018; Adeel, 2023). The researcher posited that a data quality audit was a crucial intervention in this study, aiming to enhance data quality by identifying and rectifying errors and eliminating duplicate records (Fig. 3)

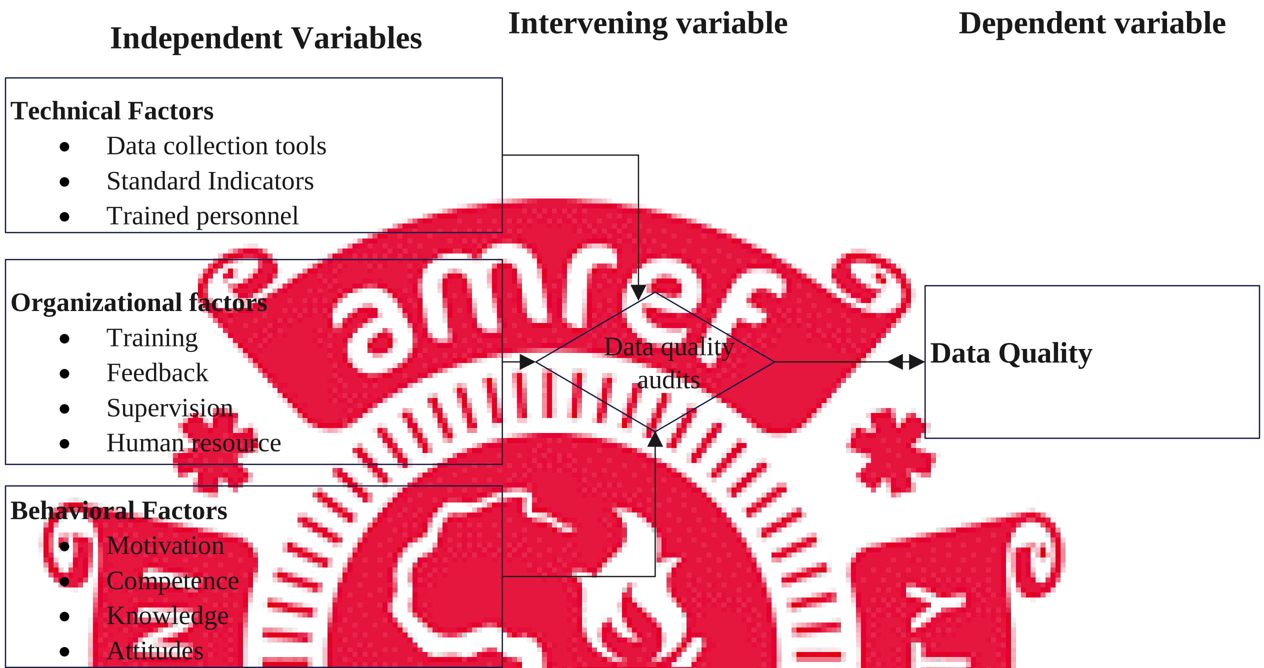

Figure 3: Conceptual framework

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

This chapter includes the study design, the study area/setting, the target population, sample and sampling procedures, data collection instruments, validity and reliability, data collection procedures, data analysis and presentation, ethical considerations, study constraints, and limitations.

3.2 Study Design

A Cross-sectional quantitative and qualitative research approach was used; the quantitative approach gathered information through scheduled interviews using the structured questionnaire, while key informant interviews helped to gather views about factors influencing data quality from the key informants. This approach allows one to get information from many people at once. The researcher preferred the cross-sectional approach because it is fast and cheaper (Kesmodel, 2018).

3.3 Study Area

The research was conducted in health facilities in Maridi County, located in Western Equatoria State. It is surrounded to the west by Ibba County, to the east by Mundri West County, and to the north by Mvolo County. It also borders the Democratic Republic of the Congo to the southwest, Lakes State (Wulu County) to the Northwest, and Yei County in Central Equatoria State to the southeast. The population of Maridi County was 82,461 in 2008. In 2020, the population had increased to 92,205 (South Sudan Bureau of Statistics, 2022). The County has five Payams Kozi, Landili, Maridi (County Seat), Mambe, and

Ngamunde. The languages spoken by the six ethnic groups include Baka, Mundu, Avukaya, Zande, Moro Kodo, and Wetu. The County has 12 functional health facilities submitting routine health information as of June 2022 (1 hospital, 5 PHCCs, and 6 PHCUs).

3.4 Why Maridi County

The researcher chose Maridi County in Western Equatoria State due to its diverse population and susceptibility to subnational violence and disease outbreaks. The county is home to reputable organizations like AMREF, AFSS, and CUAMM, which have invested heavily in strengthening the health system and training mid-level healthcare personnel. However, the routine health information system in Maridi County falls short of the national target, making it a potential study area for data-driven interventions.

3.5 Target Population

Broadly, the study focused on the population of Maridi County, with its diverse characteristics and a total population of 82,461. Specifically, the participants were drawn from 12 healthcare facilities with a total population of 146 health workers. Out of these health workers, the researcher selected participants from the staff with an assigned role in data management, including clinical officers, Nurses, Midwives, and Community Health workers. These staff are involved in data collection, including outpatient and inpatient registration (data clerks) and preparation or reviewing of the monthly reports (facility in charge).

3.5.1 Inclusion Criteria

Health workers working in functional healthcare facilities who have been employed for more than six months participated in the study because they are deemed to be very familiar with recording and compiling healthcare data.

3.5.2 Exclusion criteria

Staff whose health facilities were not operational at the time of the study and those who worked less than six months were not selected for the interview because they were considered to have little knowledge of health management information systems since the researcher was interested in generating reliable information for the study and some had not participated in managing routine health information system. Health facilities which are not operational during the data gathering period were not considered in the study because those facilities would provide no complete and accurate information and

3.6 Sample and Sampling Procedures

3.6.1 Sampling Procedure

M[illegible] County was selected through convenient sampling because it was easily accessible to the researcher, and since the 12 functional health facilities were manageable, all were considered for the study. Using probability proportional to size, health workers were chosen using simple random sampling for the quantitative study. A piece of paper with the options yes and no was cut into small pieces and folded. Each health professional was asked to choose one at random. Those who chose yes were then eligible for the interview and became research participants; the reason for using this was due to the small sample size in the health facilities, which can be easily managed using this method.

For the qualitative study, 12 Key informants who are deemed experienced with the collection of health data and familiarity with the behavioral, organizational, and technical aspects that affect data quality at various levels were selected purposively for in-depth interviews, and the interviewees were identified from multiple healthcare facilities which consist of one staff per facility and those who took part in the interview were all head of departments and the inclusion criteria were health workers who work with routine health data or those directly involved in the compilation of health facility reports.

3.7 Sample Size

The study includes each of the 12 healthcare facilities, and the Fisher Exact, a precise formula was utilized to determine the sample size of health workers. This method helps determine the optimum sample size.

$$N=\frac{Z^2PQ}{e^2}$$

Where; n is the sample size,

Z is the normal standard deviation

, p is the target population estimated at 0.5 percent, and

e is the degree of precision and applying the formulae

n= (1.96)^2*(0.5)*(0.5)/0.05^2=**384**

Since there are fewer than 10,000 people, the population correction factor (nf) was utilized.

N =n/(1+n/N)

nf=384/ (1+384/146) =**106**

Health workers were selected through probability proportional to the size of each health facility, which was calculated using **nx=x/No*n**. Where **nx** denotes a sample for a particular facility, the number of healthcare professionals in each facility is **x.**, **No** overall health workers available in the health facility, and the sample size is **n.**

Table 1: Population Sample Distribution

S/No	Health facility type		Sample size (nx)	Total population (x)
1	Maridi Hospital		43	59
2	Bethsaida PHCC		9	13
3	Don Bosco PHCC		12	16
4	Woko PHCC		7	10
5	Olo PHCC		9	12
6	Dukudu Olo PHCC		6	8
7	Kozi PHCU		7	9
8	Chochoro PHCU		5	7
9	Mabirindi PHCU		3	4
10	Longhua PHCU		3	4
11	Amaki PHCU		1	2
12	Make 2 PHCU		1	2
	Total		**n=106**	**No=146**

3.8 Data Collection Instruments

With minor modifications, the performance of routine information system management (PRISM) version 3.1 was used (Appendix 1). The tool was sorted so that questions not relevant to the study were removed and excluded to develop a pertinent tool for the study. It serves as the foundation for the questions, and the researcher anticipated that the remaining questions would add value to the tool's Validity. It was divided into the following four sections:

The first section contained inquiries into the socio-demographics of the healthcare professionals, such as their age, education, job history, and others. Sections two and three of the questionnaires were to identify behavioral, organizational, and technical elements connected to data quality. The fourth section consists of interviews with the key informants guided by a Key Informant Interview guide (Appendix 2) to collect qualitative, in-depth information/data on the departments' data quality.

3.9 Validity and Reliability

3.9.1 Validity

A validity test was conducted using Spearman's rank correlation between the questions to determine the data collection tool's accuracy in measuring the patterns of interest (Haele & Twycross, 2015). If Sig. <0.05, the question/instrument is valid, and if Sig. > 0.05, the question/instrument is not valid. However, Qn103 to Qn123 were manually sorted and reviewed because they statistically failed the validity test, and the remaining questions that failed the test were eliminated or excluded from the analysis. (Appendix 3)

3.9.2 Reliability

Using Cronbach's Alpha statistic, the instrument's Reliability was determined whereby if Cronbach's Alpha > 0.6, the instrument is reliable; otherwise, it is not if it is < 0.6. According to the data in the table below, Cronbach's Alpha was higher than 0.6, indicating that the tool was reliable.

Table 2: Reliability Test Score

Reliability Statistics		
Cronbach's Alpha		**N of Items**
Score	0.694	73

3.9.3 Dependability

Dependability is a strategy for guaranteeing that the same study conducted under similar conditions yields the same results. The researcher employed suitable qualitative data-gathering instruments and unique, in-depth interview guidelines to guarantee the study's Dependability (Appendix 2). This enables readers to evaluate the degree to which relevant research practices have been followed.

3.9.4 Transferability

The researcher conducted in-depth interviews with a wide range of participants from all the health facilities to ensure that the findings are applied to diverse patterns and identified the recurrent themes and patterns concerning the variables that impact data quality, the answers environments and represent a more comprehensive range of backgrounds (Table 6).

3.10 Data Collection Procedures

Quantitative data were gathered through face-to-face scheduled interviews with a facility in charge (data clerks, Health departments), which took place at the respective health facilities, printed questionnaires consisting of open and close-ended questions given to respondents after thoroughly explaining and consenting to participate, the respondent filed the questionnaires and submitted to the research assistants. The allocated time for the interview was 45 minutes, although most of the time spent on each interview varied from 25-35 minutes.

For Qualitative data, Face-to-face interviews were performed by the researcher and the interviewee. The participants identified as eligible and agreed to participate were invited for the interview, and primarily, the facility in charge and heads of department were the critical participants for key informant interviews. The interviewer has a face-to-face discussion with the interviewee using the interview guide, and although each in-depth interview was given a minimum of 30 minutes, the actual time spent on each interview varied from 20-25 minutes. All interviews were performed in Local Arabic and English to ensure clarity and to reduce the likelihood that the meaning of the data would be changed through translation. All information or answers for the interviews were written in a notebook by the research assistants for further analysis.

3.11 Data Analysis and Presentation

3.11.1 Data Processing

The raw data were checked for errors and completeness at the field level. The rule was to remove any questionnaire with more than 10% unanswered questions from entry. Since none of the questionnaires reached the elimination threshold, all 106 questionnaires were considered for entry. The data were then entered into SPSS version 25, cleaned, and analyzed to check for missing values (Table 4), and any missing value of less than 10% was considered negligible. Manually filled vital informant interview guides were typed into word and stored for analysis.

3.12 Data Analysis

The quantitative data were analyzed using statistical software IBM-SPSS version 25. Data on the demographic characteristics of the respondents were compiled using descriptive statistics. Tables and graphs were created using the ‘Analyze’ field in the SPSS window, and appropriate frequency distribution tables were made.

3.13 Factors Analysis

Factor analysis is a statistical technique used to uncover latent dimensions inside a dataset by analysing the correlation patterns among variables. It aids in identifying closely connected data sets and can elucidate shared patterns. Factor analysis streamlines intricate variables or objects, unveiling micro-interrelations (Gell, 2023; Tavakol & Wetzel, 2020). There are several approaches to factor analysis, including Principal Component Analysis (PCA), Exploratory Factor Analysis (EFA), and Confirmatory Factor Analysis (CFA). The researcher used the Principal Component Analysis to determine the best fit for the data. Compared to other models, Principal Component Analysis can uncover salient correlations whose outputs are easy to interpret.

3.14 Principle Component Analysis (PCA)

The principal component analysis (PCA) is a multivariate statistical technique that organizes, extracts, and groups data into components based on intercorrelations within variables (Abdi & Williams, 2010). Originating from Cauchy, it was first formulated by Karl Pearson in statistics. Hotelling later worked on the method, but it gained popularity after computers due to its complexity (Kovács et al., 2022). Principal components are linear combinations of original variables that maximize the variance of all variables, providing an approximation of the original data table using only these major components (Greenacre et al., 2022).

3.14.1 Working Model to Establish Correlations among the Variables

The working model was drawn to demonstrate the layering of the correlations among the independent variables and with the dependent variables (Fig.4)

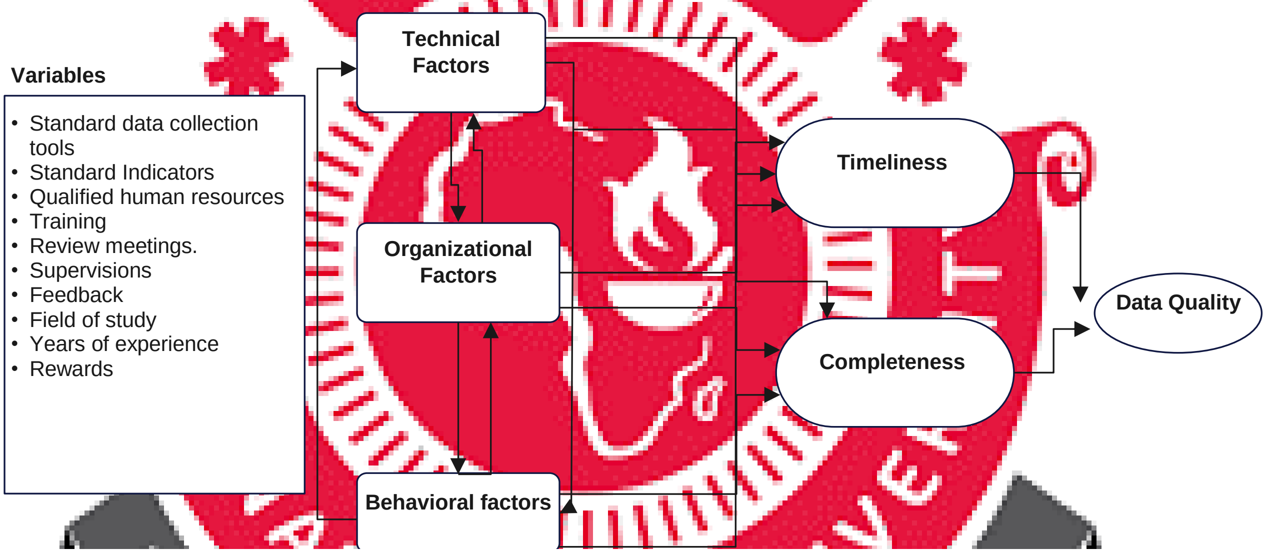

Figure 4: Working Model to Establish Correlations

Correlations among the independent variables: The independent variables were pulled together and analyzed to establish correlations.

Correlations between the independent and dependent variables: The independent variables in their groups were correlated with the two data quality variables to examine if these variables associate.

The Cohen's ranking determined the strength of the correlation. In 1988, Cohen simplified the interpretation of correlation coefficient results and recommended Pearson *r* values of 0.10-0.29, 0.30-0.49, and 0.50-1.00 to demarcate robust, moderate, and weak correlations, respectively (Hanifah et al., 2018; Gignac & Szodorai, 2016) [Table.3]

Table 3: Reference Table for the Strength of Correlation

Correlation value	**Strength**	**Color code**
0.50-1.00	Strong	
0.30-0.49	Moderate	
0.10-0.29	Weak	

3.15 Thematic Content Analysis

Thematic Content Analysis (TCA) is a descriptive method for analyzing qualitative data, including interview transcripts and other textual materials, focusing on the topic of study (Vaismoradi et al., 2013). The qualitative data was manually examined by themes, reviewed twice for accuracy and consistency, and then analyzed using a thematic framework (Table 4). The data collected were categorized into four subthemes. Data quality was one of the major themes in addition to Organizational, technical, and behavioral factors. The detailed notes of the key points were then aligned with the research objectives and coded (Appendix 6)

Table 4: Thematic Framework

Factors influencing data Quality in routine health information system			
Behavioral factors	**Organizational factors**	**Technical factors**	**Data quality practices**
1 Lack of motivation 2 Negative attitude towards work 3 Work overload. 4 Lack of cooperation among staff 5 Recruiting more staff to the facilities	1 Inadequate human resource 2 poor supportive supervision 3 No training of staff on HMIS 4 Lack of performance feedback to facilities	1 Lack of data collection and reporting tools 2 Lack of trained staff on HMIS	1 Data collected from facility registers and entered into monthly reports. 2 Yes, such as data verification by the person in charge. 3 Crosschecking the reports before submission 4 Prepare reports jointly with the team to avoid errors. 5 Documenting all information about the patients

3.16 Ethical Considerations

The Amref International University approval letter was presented to South Sudan's Ministry of Health, Research, and Ethics Review Board for approval (Appendix 4). The approval letter for the research and ethics board was presented at the county level to the County Health Director-Maridi for further approval (Appendix 5). Before administering the questionnaire, the participants were given consent forms after explicitly explaining what the study entails (Appendix 1). The instruments used to gather data had no names but rather codes. Keeping the questionnaire anonymous protects the identity of the respondents.

3.17 Study Constraints and Limitations

This study was conducted at rural public healthcare facilities in Maridi County, with a small geographic coverage. It only covered 12 of the 23 health facilities, representing 52% coverage. Readers should know that these results may not directly apply to urban and private healthcare facilities.

A limitation on context-specific literature emerged as a constraint because of the generality of information. The researcher expanded the literature search to the neighboring countries to make inferences on the assumption that these countries share some related characteristics. However, the complex context emerging from different health systems structures means that some readers may question some associations with practices of other countries.

The study looked at two dimensions of data quality (Completeness & timeliness) ignoring other dimensions vital in determining data quality.

CHAPTER 4: RESULTS

4.1 Introduction

The results chapter presents the results from the quantitative analysis and the results from the qualitative data analysis describing factors influencing data quality.

4.2 Presentation of the results in line with the specific objectives

4.2.1 Response rate

There was a 100% response rate, as shown below in the table (Tab 5)

Table 5: Response rate

Sample size	**Respondents**	**Responses rate**
106	106	100%

4.2.2 Key Sociodemographic Characteristics

The study involved 106 health workers in data management, 40.6% from hospitals, [illegible]% from PHCCs, and 18.9% from primary health care units. A significant proportion of participants were male, accounting for 74.5%. Most had certificates and diplomas, surpassing other qualifications. The average age range was 26-40 years (Table 6).

Table 6: Key Sociodemographic Characteristics

Variable	Frequency	Percent
Age		
less than 25 years	8	7.5
26-40 years	63	59.4
41-56 years	31	29.2
Above 56 years	4	3.8
Total	106	100
Gender		
Female	27	25.5
Male	79	74.5
Total	106	100
Level of Education attained		
None	7	6.6
Certificate	72	67.9
Diploma	25	23.6
Bachelor’s degree	2	1.9
Total	106	100

4.2.3 Data Quality in Maridi County

The completeness and timeliness of the reports submitted to the County health department were utilized to assess the quality of health care data in Maridi County.

4.2.3.1 Timeliness.

The standard practice in South Sudan is that reports for the previous month are submitted to the County Health Department by the fifth of the following month. The results, however, show that only eight health facilities meet this timeline, translating to 67% performance.

4.2.3.2 Completeness.

It is required that all reports be submitted to the County Health Department for the reporting period. The overall performance fell short of the requirements for completeness as 9 of 12 health facilities submitted all reports to the County Health Department, a performance of 75%.

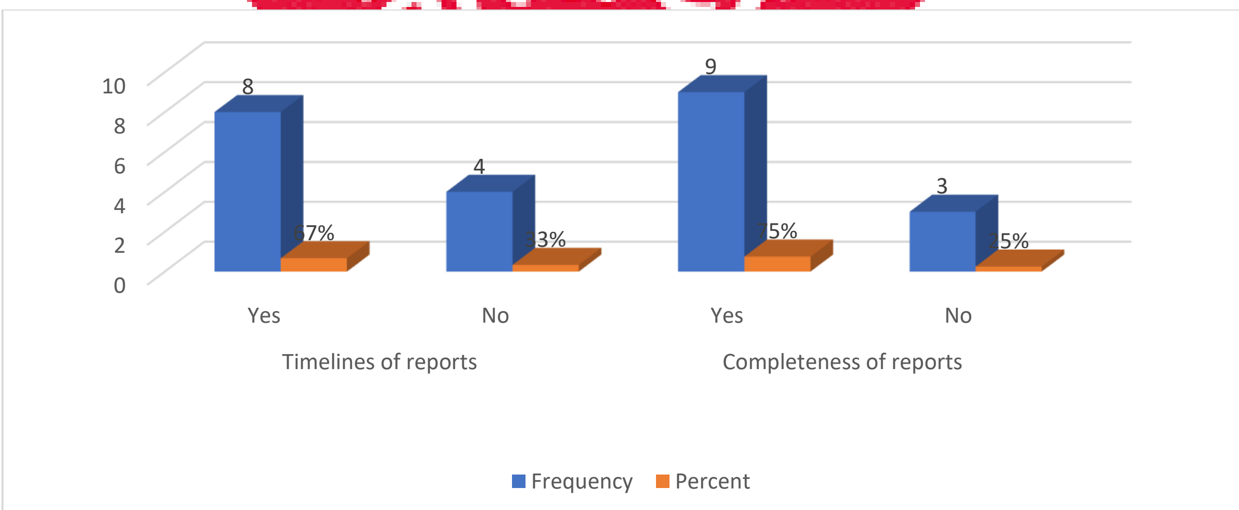

Figure 5: Timeliness and Completeness

Descriptive Statistics for the factors influencing data quality in Maridi County, South Sudan.

4.2.4.1 Technical Factors.

The majority of the respondents, 94 (88.7%), agreed that their healthcare facilities use standardized data collection methods. Furthermore, 66 (62.3%) stated that they receive data collection tools tailored to their specific requirements on a regular basis. Additionally, 86 (81.1%) reported that their facilities possess a collection of standardized and clearly defined indicators. Likewise, 82 (77.4%) indicated that they have staff members who have the required training to complete the requisite documents. Finally, 100 respondents (94.3%) indicated that the reporting formats are user-friendly and understandable (Table 7).

Table 7: Technical Factors

Variables	Categories	Frequency	Per cent
Availability of standard data collection tools	Yes	94	88.7
	No	12	11.3
How often are you supplied with data collection tools	depending on need	66	62.3
	Monthly	26	24.5
	Quarterly	9	8.5
	Others	5	4.7
Do you have a set of indicators that are standardized and defined	Yes	86	81.1
	No	20	18.9
Do you believe that the report and registration forms are user-friendly and simple to understand	Yes	100	94.3
	No	6	5.7
Do you have qualified human resources who can fill out formats?	Yes	82	77.4
	No	24	22.6

***4.2.3.2** Organizational Factors.*

Above half of the participants (62.3%) did not receive refresher training in the past six months. Additionally, 76.4% of the respondents acknowledged the occurrence of review meetings, while 80.2% received supportive supervision. Furthermore, 80.2% of the respondents acknowledged receiving data quality supervisions from the County Health Department. During these supervisions, 81.1% of the respondents revealed that their supervisors evaluated the accuracy of the data. Lastly, 57.5% of the respondents received performance feedback from the county health department (Figure 6).

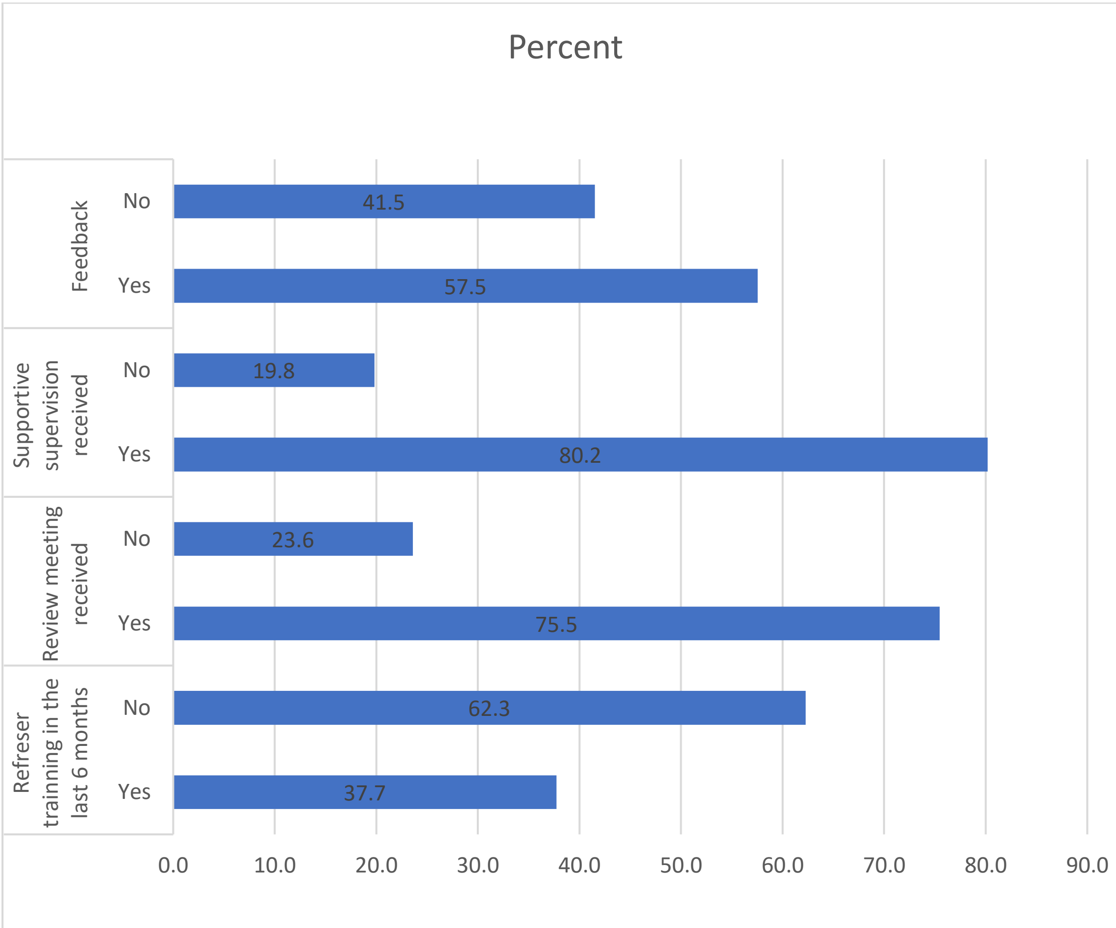

Figure 6: Organizational factors

4.2.4.3 Behavioural Factors.

Only 38.7% of the respondents had additional certifications, such as Lab technician and Public Health. This was followed by 25.5% of the respondents who had certifications in nursing. Community health workers made up 16% of the respondents, while clinical medicine and midwifery accounted for 11.3% and 8.5% respectively. With respect to years of experience, 44.3% of individuals had 1-6 years of work experience, while 34.9% were those with more than 12 years of experience. 69.8% of the health workers expressed motivation to do HMIS activities, while 85.8% of the respondents actively participated in the gathering or aggregation of data within the health facility (Table 8).

Table 8: Behavioural Factors

Variables	Categories	Frequency	Per cent
Respondent education's field of study	CHW	17	16.0
	Nurse	27	25.5
	Midwife	9	8.5
	clinical medicines	12	11.3
	Others	41	38.7
Years of work experience at work	1-6 years	47	44.3
	7-12 years	22	20.8
	Above 12 years	37	34.9
Are there any rewards or motivations for the HMIS process	Yes	74	69.8
	No	32	30.2
If yes, what types of rewards	Cash	31	29.2
	Training	26	24.5
	Appreciation	14	13.2
	Others	3	2.8
Have you taken part in the collection or aggregation of data from the health facility	Yes	91	85.8
	No	15	14.2

4.2.5 Inferential Statistics for Factors Influencing Data Quality in Maridi County, South Sudan.

4.2.5.1 Correlations among the Independent Variables.

Factor analysis was done to uncover deeper interactions among the independent variables to demonstrate their interdependence as a system. Over half of the variables show correlations from weak to strong. Years of staff work experiences has moderate significant correlation with respondents' field of study (r=-0.305, P=0.001<0.05) and weak significant correlations with qualified human resources to fill the formats and taking part in the collection and aggregation of data from the facility (r=-0.172, P=0.039<0.05) & (r=-0.171, P=0.04<0.05) respectively.

There was moderate significant correlation between motivations of facility staff with review meetings conducted on data quality (r=0.409, P=0.000<0.05) and a set of standardized indicators (r=0.366, P=0.000<0.05) whereas training on HMIS (r=0.215, P=0.013<0.05), supportive supervisions on data quality (r=0.240, P=0.007<0.05), regular feedback from the county health department (r=0.284, P=0.002<0.05), qualified human resources (r=0.184, P=0.029<0.05) and taking part in the collection of data (r=0.264, P=0.003<0.05) has significant weak correlations with motivations of facility staff.

There was weak significant correlation between review meetings conducted (r=0.295, P=0.001<0.05), supportive supervisions on data quality (r=0.289, P=0.001<0.05), availability of standard data collection tools (r=0.278, P=0.002<0.05), set of standardized indicators (r=0.226, P=0.01<0.05), user-friendly and simple forms (r=[illegible]99, P=0.025<0.05) and qualified human resources (r=0.189, P=0.026<0.05) with training on HMIS whereas taking part in data collection and aggregation has moderate significant correlations with training on HMIS (r=0.316, p=0.000<0.05).

Review meetings has moderate significant correlations with supportive supervisions (r=0.449, P=0.000<0.05), set of standard indicators (r=0.470, P=0.000<0.05) and availability of standard data collection tools (r=0.363, P=0.00<00.05) whereas weak moderate significant correlations were noted between review meetings and regular feedback ($r=0.299$, $P=0.001<0.05$), qualified human resources ($r=0.177$, $P=0.034<0.05$), taking part in data collection and aggregation ($r=0.221$, $P=0.011<0.05$).

Supportive supervisions have significant moderate correlations with regular feedback ($r=0.302$, $P=0.001<0.05$), availability of standard data collection tools ($r=0.345$, $P=0.000<0.05$), set of standard indicators ($r=0.365$, $P=0.000<0.05$) and weak significant correlations with user-friendly and simple forms ($r=0.186$, $P=0.028<0.05$).

The correlation between regular feedback and availability of standard data collection tools ($r=0.303$, $P=0.001<0.05$), set of standard indicators ($r=0.475$, $P=0.000<0.05$) was significantly moderate.

Availability of standard data collection tools has significant strong correlation with set of standard indicators ($r=0.513$, $P=0.000<0.05$) and significantly moderate correlations with user-friendly and simple forms ($r=0.428$, $P=0.000<0.05$) and qualified human resources ($r=0.305$, $P=0.001<0.05$).

Set of standard indicators significantly correlated weak with user-friendly and uncomplicated forms ($r=0.195$, $P=0.023<0.05$), qualified human resources ($r=0.200$, $P=0.02<0.05$).

Finally, user-friendly and straightforward forms and qualified human resources ($r=0.355$, $P=0.000<0.05$) significantly correlated moderately (Table 9).

Table 9: Correlations among independent variables

Correlation Matrix												
	Years of work experiences	motivations at work place	Respondent education's field of study	Trained on HMIS	Review meetings conducted	supportive supervisions	Regular feedback	standard data collection tools available	A set of standardized indicators	user-friendly and simple forms	qualified human resources	Taken part in the collection of data
Years of work experiences	1.000	0.000	-0.305	-0.083	-0.091	-0.054	0.025	0.038	-0.030	0.118	-0.172	-0.171
P-Values		0.498	0.001	0.199	0.176	0.291	0.400	0.349	0.379	0.113	0.039	0.040
Motivations of staff	0.000	1.000	-0.065	0.215	0.409	0.240	0.280	0.284	0.366	0.017	0.184	0.264
P-Values	0.498		0.254	0.013	0.000	0.007	0.002	0.002	0.000	0.432	0.029	0.003
Respondent field of study	-0.305	-0.065	1.000	-0.069	0.031	-0.084	0.028	-0.071	0.012	-0.023	0.266	0.092
P-Values	0.001	0.254		0.242	0.375	0.197	0.387	0.235	0.452	0.409	0.003	0.174
Trained on HMIS	-0.083	0.215	-0.069	1.000	0.295	0.289	0.142	0.278	0.226	0.191	0.189	0.316
P-Values	0.199	0.013	0.242		0.001	0.001	0.073	0.002	0.010	0.025	0.026	0.000
Review meetings	-0.091	0.409	0.031	0.295	1.000	0.449	0.299	0.363	0.470	0.152	0.177	0.221
P-Values	0.176	0.000	0.375	0.001		0.000	0.001	0.000	0.000	0.059	0.034	0.011
supportive supervisions	-0.054	0.240	-0.084	0.289	0.449	1.000	0.302	0.345	0.365	0.186	0.127	0.002
P-Values	0.291	0.007	0.197	0.001	0.000		0.001	0.000	0.000	0.028	0.097	0.492
Regular feedback from the CHD	0.025	0.280	0.028	0.142	0.299	0.302	1.000	0.303	0.475	0.125	0.093	0.097
P-Values	0.400	0.002	0.387	0.073	0.001	0.001		0.001	0.000	0.101	0.171	0.160
Availability of standard data collection tools	0.038	0.284	-0.071	0.278	0.363	0.345	0.303	1.000	0.513	0.428	0.305	-0.060
P-Values	0.349	0.002	0.235	0.002	0.000	0.000	0.001		0.000	0.000	0.001	0.272
A set of standardized indicators	-0.030	0.366	0.012	0.226	0.470	0.365	0.475	0.513	1.000	0.195	0.200	0.012
P-Values	0.379	0.000	0.452	0.010	0.000	0.000	0.000	0.000		0.023	0.020	0.452
user-friendly and simple forms	0.118	0.017	-0.023	0.191	0.152	0.186	0.125	0.428	0.195	1.000	0.355	-0.099
P-Values	0.113	0.432	0.409	0.025	0.059	0.028	0.101	0.000	0.023		0.000	0.155
qualified human resources	-0.172	0.184	0.266	0.189	0.177	0.127	0.093	0.305	0.200	0.355	1.000	-0.026
P-Values	0.039	0.029	0.003	0.026	0.034	0.097	0.171	0.001	0.020	0.000		0.397
Taken part in the collection or aggregation of data	-0.171	0.264	0.092	0.316	0.221	0.002	0.097	-0.060	0.012	-0.099	-0.026	1.000
P-Values	0.040	0.003	0.174	0.000	0.011	0.492	0.160	0.272	0.452	0.155	0.397	

4.2.5.2 Correlation between Technical Factors and Data Quality.

All correlations between the technical factors and the data quality variables are insignificantly weak. The correlation of the set of standard indicators and timeliness is insignificantly higher amongst the lower values at (r=-0.213, p=0.253>0.05), followed by how often supplied with data collection tools and completeness at (r=-0.204, p=0.262>0.05) shows weak, insignificant correlations. The correlation between a set of standard indicators and completeness was slightly higher at (r=-0.174, p=0.294>0.05) than the correlation between the frequency of supply of data collection tools and timeliness at (r=-0.125, p=0.349>0.05). Availability of qualified human resources showed no correlation with timeliness at (r=0.00, p=0.500>0.05) and a much weaker insignificant correlation with completeness at (r=0.11, p=0.366>0.05), respectively (Table 10).

Table 10: Correlations between Technical Factors and Data Quality

Correlation Matrix			
	How often supplied with data collection tools	A set of standard indicators	Qualified human resources
All the monthly RHIS submitted to the CHD (Completeness)	-0.204	-0.174	0.111
P-Values	0.262	0.294	0.366
Are monthly RHIS reports submitted on time (Timeliness)	-0.125	-0.213	0.000
P-Values	0.349	0.253	0.500

*. Correlation is significant at the 0.05 level (2-tailed).

The qualitative data also revealed the impact of the consistency of data collection tool provision on reporting rates; some participants identified the absence of data collection and reporting tools at the health facility, as well as the lack of qualified personnel in the health information management system, as factors that negatively affect data quality.

These qualitative outputs supported the quantitative data that emphasized the connections between the three factors and the two data quality measures. Among these, four exhibited negative correlations.

"Sometimes, if data collection tools like registers and reporting forms are not supplied regularly, our report is affected because the patients seen during that period are not going to be registered or their information will not be documented'' (Key informant 4)

Table 11: Correlation between Organizational Factors and Data Quality

Correlation Matrix				
	Regular feedback from the CHD	Supportive supervisions on data quality	Training on HMIS	Review meetings conducted to discuss data quality
All the monthly RHIS submitted to the CHD (Completeness)	0.683	0.258	0.488	0.522
P-Values	0.007	0.209	0.054	0.041
Are monthly RHIS reports submitted on time (Timeliness)	0.120	0.158	0.239	0.426
P-Values	0.356	0.312	0.227	0.083

*. Correlation is significant at the 0.05 level (2-tailed).

The results show strong significant correlations with regular feedback from the CHD and completeness of reporting (r=0.683, p=0.007<0.05) and review meetings (r=0.522, p=0.041<0.05). Moderate insignificant correlations resulted from the training of staff on HMIS and completeness (r=0.488, p=0.054>0.05), review meetings, and timeliness (r=0.426, p=0.083>0.05). The associations between training of staff on HMIS and timeliness, data quality supervision and timeliness and completeness, and regular feedback with timeliness were insignificantly weak (r=0.239, p=0.227>0.05), (r=0.158, p=0.312>0.05), (r=0.258, p=0.209>0.05) and (r=0.120, p=0.356>0.05).

The findings revealed that regular feedback from the County Health Department to the health facilities staff and review meetings conducted to discuss data quality has a strong significant correlation with data quality.

The qualitative results corroborated the quantitative findings, as most participants considered that the organizational factors influencing that quality were the lack of human resources to perform the HMIS work, poor supportive supervision by the supervisors, lack of performance feedback to help staff know how their health facility is performing, and inadequate staff training on the tools.

"Only the in charge and data clerk were trained in this facility for the new HMIS tools, but the rest of the department's heads were not trained, yet they are required to use these tools." (key informant 1)

"I am the only clinical officer clerking patients in this facility. If the nurse is not present, I occasionally have to do ward rounds and even dispense drugs, which can get tiresome." (Key informant 5).

Table 12: Correlation between Behavioural Factors and Data Quality

Correlation Matrix			
	Respondent education's field of study	Years of work experiences at work	Rewards or motivations
All the monthly RHIS submitted to the CHD (Completeness)	-0.113	-0.555	-0.098
P-Values	0.363	0.031	0.381
Are monthly RHIS reports submitted on time (Timeliness)	-0.277	0.062	0.478
P-Values	0.191	0.424	0.058

Behavioral factors and data quality variables correlated to some extent. Years of experience have significant solid correlations with completeness (r=-0.555, p=0.031<0.05), while motivations showed an insignificant moderate relationship with timeliness (r=0.478, p=0.058). The rest of the variables showed a weaker insignificant correlation with completeness and timeliness at (r=-0.277, p=0.191>0.05) (r=0.061, p=0.424>0.05) (r=-0.111, p=0.363>0.05), (=-0.09759, p=0.381>0.05) respectively.

Years of work experience have a significant correlation with data quality; however, the rest of the factors have insignificant correlations.

On a qualitative note, these results are consistent with the perspectives generated during the key informant interviews, where most participants thought lack of motivation to staff performing health management information tasks such as incentives, appreciation, and negative attitude towards work by some of the staff, work overload and lack of cooperation among the staff affected data quality.

"Facility staff sometimes have negative attitudes toward their jobs, and this is largely a result of the low pay some staff members give as an excuse for cultivating before coming to work so that they can support their families" (Key informant 3)

"I have much work to do here. We have many registers to fill out, and when it comes to reports, I have to gather all the reports from the wards'' (key informant 7)

''Lack of cooperation among us sometimes affects the quality of data because some staff disappear in the facility during a period of reporting'' (key informant 2)

CHAPTER 5: DISCUSSIONS

5.1 Introduction

This section provided an opportunity for the researcher to comprehend the findings in relation to the study's objectives. It discussed the findings' theoretical relevance and elaborated on their scientific analogy with existing literature.

5.2 Response Rate

This was a scheduled interview, so the researcher made prior arrangements with the respondent and agreed on a convenient time. The researcher's ease of access to all healthcare facilities made the 100% study response rate (Table 5). In contrast to afternoon hours, when some staff members might have left for lunch and been unable to return for afternoon duty, the researcher used the morning hours of 9:00 a.m. to 12:00 p.m. when most health workers are present at their workplace and followed up with participants who were not present during the visits. No participant who met the study's eligibility requirements declined to participate in the interview.

5.3 Data Quality in Maridi County, South Sudan

The quality of data was assessed by the timeliness and completeness of the data submitted to the County Health Department. The performance, however, fell short of the national requirements of 90% for completeness and 85% for timeliness. The low performance demonstrates the chronic challenges affecting the performance of data quality. Notably, the digitalization of the RHMIS requires that reporting facilities be connected to the internet so that data is electronically transmitted from the health facility to the County health Department. Unfortunately, all the health facilities in Maridi County use a manual approach, thus limiting the efficiency of transmitting data in real time. Many African Countries, including those that have shown significant investment in health systems strengthening also experience challenges in meeting timeliness and completeness targets. In Uganda, the national average reporting timeliness and completeness from 2020-2021 staggered between 44% and 70%, below the national targets of 90% (Nansikombi et al., 2023). A study in Ethiopia found that health centers in West Gojjam Zone have a data quality of 74%, below national targets due to complex health information systems and inadequate problem-solving skills (Chekol et al., 2023).

5.4 Factors influencing -Data Quality in Maridi County, South Sudan

5.4.1 Technical Factors and Data Quality

The results revealed a remarkable depiction of the technical components necessary to operate a regular health information framework effectively. The presence of standardized data collecting instruments in 62.3% of health facilities and standard indicators in [illegible]% of health facilities indicates a significant disparity. This is reinforced by the little link between the technical aspects and the data quality. The inclusion of one technical indication over others does not enhance the system's performance. The qualitative findings further supported the necessity of ensuring consistent dissemination of all tools and technical protocols throughout all healthcare institutions, encompassing competent personnel, standardized data collection instruments, and user-friendly reports and registration forms that are easily comprehensible. The crucial aspect of this element is the necessity for a consistent and enough provision of the technical prerequisites to enhance the effectiveness of the regular health management information framework. This study aligns with previous research conducted by Wude et al. (2020) and Wandera et al. (2019), which found that the presence of trained personnel and a standardized set of indicators strongly influence data quality. It is also consistent with the findings of Mucee et al. (2016) in Tharaka Nithi County, which demonstrated that the competence of staff and the use of standardized data collection tools have an impact on data quality. Hlaing et al. (2022) also identified multiple reporting, inexperienced personnel, and a lack of reporting tools as key factors affecting data quality.

5.4.2 Organizational Factors and Data Quality

On a related point, the organizational factors varied, with certain indicators, such as refresher training and regular feedback, having a greater positive impact on data quality during quantitative analysis. However, the qualitative findings revealed that insufficient human resources and inadequate supportive supervision were identified as factors that negatively affect data quality. The results align with prior research conducted by Glette et al. (2021), which demonstrated that organizational elements such as people training, ongoing feedback on data quality, supportive supervision, and favorable working circumstances for healthcare staff are influential. According to Lemma et al. (2020), implementing capacity-building strategies, such as training, data quality evaluation, and feedback supply to healthcare facilities, can enhance the quality of data.

Moloko et al. (2022) conducted prior investigations in Tshwane, South Africa, which revealed that training, supportive supervision, and adequate human resources exert an impact on the data's quality. These findings support the research conducted in Northwest Ethiopia by Afework (2022), which identified deficient feedback systems, insufficient human resources, and inadequate training as barriers to data quality. This aligns with a study conducted in Kenya by Cheburet et al. (2016), which found that support supervision positively impacted the data quality. Similarly, a survey conducted by Shiferaw et al. (2017) in Gojjamzone, Northeast Ethiopia, demonstrated that supportive supervision, HMIS training, and providing feedback to health facilities were significantly associated with data quality. These results align with the findings of Tulu et al. (2021) in Ethiopia, which revealed that supportive supervision and HMIS training were significantly associated with data quality. Thus, the findings of prior studies align with the current findings of this study, indicating that it is crucial to address these organizational issues to enhance data quality.

5.4.3 Behavioral Factors and Data Quality

The study found that education level and years of experience significantly impact data quality. The qualitative findings revealed that factors such as lack of motivation among staff responsible for health management information tasks, negative attitude towards work, work overload, and lack of cooperation among staff also influence data quality. These findings align with the research conducted by Kleiman et al. (2020) and Chanyalew et al. (2021), which indicate that employee competence, motivation, and attitude toward health information systems are behavioral factors that impact data quality. These findings are also consistent with the study by Glèlè Ahanhanzo et al. (2014), which identified worker demotivation and low capability as factors contributing to inadequate data quality. According to Hlaing et al. (2022), work burdens influence the quality of healthcare data owing to a lack of available personnel, which can lead to excessive workloads. This study corroborates the hypothesis that the proficiency of healthcare professionals, as assessed by their level of education and engagement in their work, is connected to the accuracy of the data. Furthermore, all these findings align with the current study.

In a separate study conducted by Moses et al. (2019) in South Sudan, the effectiveness of gathering, analyzing, and understanding data is hindered by a lack of skilled workers in healthcare facilities. Insufficiently trained individuals may be unable to collect precise or inaccurate data, lowering the overall quality of routine health information. In contrast, Haftu et al. (2021) found that lacking a competent HMIS focal person and lacking motivation for HMIS responsibilities are hindrances to ensuring data accuracy in Ethiopia. These findings align with the current study's findings, suggesting that addressing all behavioral factors will guarantee data quality.

5.4.4 The Interdependence of the Factors Influencing Data Quality on each other

Although this was a side purpose, it was necessary to shed light on the interactions among the independent variables to demonstrate the holistic effect of these factors on the performance of the RHMIS in Maridi County, South Sudan. It was also essential to demonstrate the theoretical fit of the study, being grounded in the Systems & Systems Evaluation theory. The results showed reasonably strong to moderate interactions among the independent variables. Overall, 66% of the independent variables correlated with each other, 2% strongly, 29% moderately, and 36% weakly. These correlations showed a mix of relationships through all three categories of technical, organizational, and behavioral factors. Such interaction, therefore, showed the systems nature of the factors that impact the regular health information systems functional and agreed with the systems theory.

In 1998, Ken Orr published a data quality and systems theory book. In his narration, information systems are embedded in a circle of real-world feedback control systems. The book reinforced a system of thought on data quality and advised a goal-centered and organizational approach (Orr, 1998).

CHAPTER 6: SUMMARY, CONCLUSION AND RECOMMENDATIONS

6.1 Introduction

This chapter examines the study's conclusions, recommendations, and suggestions for future research.

6.2 Conclusions

The study reveals challenges in data quality in RHMIS in Maridi County, such as inadequate training and lack of reporting tools as technical factors, while organizational factors include inadequate resources, poor supervision, and regular feedback, whereas lack of motivation, work overload, and attitude towards work are behavioral factors. It suggests that data quality can be achieved when requirements are evenly distributed.

6.3 Recommendations

1. The County Health Department to conduct refresher training to all staff working in the health facilities on HMIS, this will enable them to use the instruments for gathering and reporting data efficiently and effectively.
2. The County Health Department should ensure the availability of the instruments for gathering and reporting data by ensure adequate supply of the tools in all the health facilities to avoid issues of stock outs.
3. Addressing the human resource gap by recruiting enough staff in the health facility by County health department to close the gap and this can address the challenge of work overload since staff will be enough to perform their duties in the health facilities.

4. Provide supportive supervision by the County Health Department, the State Ministry of Health, and implementing partners, which should frequently be to the health facilities, and mentorship should be conducted during those visits.
5. Provide regular feedback on data quality by the County Health Department to the health facilities, and this will help them understand their performance status, hence encouraging them to work hard.
6. The County Health Department should motivate staff in the health facilities, especially those performing HMIS data, through appreciation or incentives.
7. Health facility staff should be encouraged to have a positive attitude toward their work and cooperate with each other since good attitudes and cooperation among staff considerably contribute to the quality of data in the health facility. Addressing this can be a huge success.
8. Standardization of tools and indicators by the Ministry of Health which can be easily understood by the health workers

6.4 Suggestions for Further Research

Future research should consider increasing the sample size and evaluating the data's quality at each facility level.

The study also examined public health facilities while omitting private ones; as a result, further research should be considered to include private health facilities.

REFERENCES

Abdi, H., & Williams, L. J. (2010, July). Principal component analysis. *WIREs Computational Statistics, 2*(4), 433–459. https://doi.org/10.1002/wics.101

Adane, A., Adege, T. M., Ahmed, M. M., Anteneh, H. A., Ayalew, E. S., Berhanu, D., … & Lemma, S. (2021). Exploring data quality and use of the routine health information system in Ethiopia: A mixed-methods study. *BMJ Open*, 11(12), e050356.

Adeel (2023). *Mediating & Intervening Variables: Overview & Examples*. https://study.com/learn/lesson/mediating-intervening-variables-overview-examples.html

Adejumo, A. (2017). An assessment of data quality in routine health information systems in Oyo State, Nigeria.

Afework, C. (2022). *Data quality and associated factors in the routine health information system among health centers of West Gojjam Zone, Northwest Ethiopia,* 2021 (Doctoral dissertation).

Aqil, A., Lippeveld, T., & Hozumi, D. (2009). PRISM framework: A paradigm shift for designing, strengthening and evaluating routine health information systems. *Health Policy and Planning, 24*(3), 217-228.

Bhandari (2023). *Independent vs. Dependent Variables | Definition & Examples*. https://www.scribbr.com/methodology/independent-and-dependent-variables/

CDC (2020). *Key characteristics of data quality in public health surveillance*. https://www.cdc.gov/ncbddd/birthdefects/surveillancemanual/chapters/chapter-7/chapter7.5.html

Chanyalew, M. A., Yitayal, M., Atnafu, A., & Tilahun, B. (2021). Routine health information system utilization for evidence-based decision making in Amhara national regional state, Northwest Ethiopia: A multi-level analysis. *BMC Medical Informatics and Decision Making*, *21(*1). https://doi.org/10.1186/s12911-021-01400-5

Cheburet, S. K., & Odhiambo-Otieno, G. W. (2016). Organizational factors affecting data quality of routine health management information system quality: Case of Uasin Gishu County Referral Hospital, Kenya. *International Research Journal of Public and Environmental Health*, 3(9), 201-208. http://dx.doi.org/10.15739/irjpeh.16.026

Chekol, A., Ketemaw, A., Endale, A., Aschale, A., Endalew, B., & Asemahagn, M. A. (2023). Data quality and associated factors of routine health information system among health centers of West Gojjam Zone, northwest Ethiopia, 2021. *Frontiers in Health Services*, *3*, 1059611.

Cherry (2022). *What is a dependent variable?* https://www.verywellmind.com/what-is-a-dependent-variable-2795099

Dagnew, E., Woreta, S. A., & Shiferaw, A. M. (2018). 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*, *18*, 1-8.

Epizitone, A., Moyane, S. P., & Agbehadji, I. E. (2023, March). A systematic literature review of health information systems for healthcare. In Healthcare (Vol. 11, No. 7, p. 959). MDPI.

Gell, T. (2023). *Factor analysis: Definition, types, and examples*. https://www.driveresearch.com/market-research-company-blog/factor-analysis-definition-types-and-examples/

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74-78.

Glèlè Ahanhanzo, Y., Ouedraogo, L. T., Kpozèhouen, A., Coppieters, Y., Makoutodé, M., & Wilmet-Dramaix, M. (2014). Factors associated with data quality in the routine health information system of Benin. *Archives of Public Health*, 72(1), 1-8.

Glèlè Ahanhanzo, Y., Ouendo, E. M., Kpozèhouen, A., Levêque, A., Makoutodé, M., & Dramaix-Wilmet, M. (2015). Data quality assessment in the routine health information system: An application of the Lot Quality Assurance Sampling in Benin. *Health Policy and Planning*, *30*(7), 837-843.

Glette, M. K., & Wiig, S. (2021, May 7). The role of organizational factors in how efficiency-thoroughness trade-offs potentially affect clinical quality dimensions: A review of the literature. *International Journal of Health Governance*, *26*(3), 250–265. https://doi.org/10.1108/ijhg-12-2020-0134

Greenacre, M., Groenen, P.J.F., Hastie, T. et al. Principal component analysis. *Nature Reviews Methods Primers* 2, 100 (2022). https://doi.org/10.1038/s43586-022-00184-w

Haftu, B., Taye, G., Ayele, W., Habtamu, T., & Biruk, E. (2021). 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*, 35(1), 15-24. https://www.ajol.info/index.php/ejhd/article/view/210746

Hanifah, A. N., Mulyana, A., & Nuraini, H. (2018). The correlation between students' reading motivation and their reading comprehension at Tenth grade students of Sman Kkota Tangerang in academic year 2018/2019. *JIPIS*, *26*(2), 6-12.

Harikumar, S. (2012). *Evaluation of Health Management Information Systems: A study of HMIS in Kerala* (Doctoral dissertation, SCTIMST).

Haug, A., Zachariassen, F., & Van Liempd, D. (2011). The costs of poor data quality. *Journal of Industrial Engineering and Management (JIEM*), 4(2), 168-193.

Hlaing, T., & Zaw M. M. (2022). Factors affecting data quality of health management information system at township level, Bago Region, Myanmar. *International Journal of Community Medicine and Public Health* *9*(3), 1298. Doi: 10.18203/2394-6040.ijcmph20220686.

Hooker, C. (2011). Introduction to Philosophy of Complex Systems: A: Part A: Towards a framework for complex systems. In *Philosophy of Complex Systems* (pp. 3-90). North-Holland.

Jada, S. R., Dusabimana, A., Abd-Elfarag, G., Okaro, S., Brusselaers, N., Carter, J. Y., … & Colebunders, R. (2022). The prevalence of onchocerciasis-associated Epilepsy in Mundri West and East Counties, South Sudan: A door-to-door survey. *Pathogens*, 11(4), 396.

Kawakyu, N., Coe, M., Wagenaar, B. H., Sherr, K., & Gimbel, S. (2023). Refining the Performance of Routine Information System Management (PRISM) framework for data use at the local level: An integrative review. *PloS One*, 18(6), e0287635.

Kesmodel, U. S. (2018). Cross-sectional studies–what are they good for?. *Acta obstetricia et gynecologica Scandinavica*, *97*(4), 388-393.

Kirimi, N. S. (2017). *Factors influencing performance of routine health information system: The case of Garissa Sub-county, Kenya* (Doctoral dissertation, University of Nairobi).

Kleiman, F., Meijer, S., & Janssen, M. (2020). Behavioural factors influencing the opening of government data by civil servants: Initial findings from the literature. *In Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance* (pp. 529-534).

Kovács, T. Z., Bittner, B., Huzsvai, L., & Nábrádi, A. (2022). Convergence and the Matthew Effect in the European Union Based on the DESI Index. *Mathematics*, 10(4), 613. https://doi.org/10.3390/math10040613

Kuyo, R. O., Muiruri, L., & Njuguna, S. (2018). Organizational factors influencing the adoption of the district health information system 2 in Uasin Gishu County, Kenya. *International Journal of Medical Research & Health Sciences*, 7(10), 48-57.

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

Li, M., Brodsky, I., & Geers, E. (2018). Barriers to use of health data in low-and middle-income countries: A review of the literature. *MEASURE Evaluation*.

Maïga, A., Jiwani, S. S., Mutua, M. K., Porth, T. A., Taylor, C. M., Asiki, G., ... & Boerma, T. (2019). Generating statistics from health facility data: the state of routine health information systems in eastern and southern Africa. *BMJ Global Health*, *4*(5), e001849.

Mathewos, T. (2015). *Community health management information system Performance and factors associated with at health post of Gurage zone, SNNPR, Ethiopia*. University of Gondar and Addis Continental Institute of Public Health.

MEASURE Evaluation. (2019). Performance of Routine Information System Management (PRISM) toolkit: PRISM tools. Chapel Hill, NC, USA: *MEASURE Evaluation*, University of North Carolina.

Meghani, A., Tripathi, A. B., Bilal, H., Gupta, S., Prakash, R., Namasivayam, V., ... & Ramesh, B. M. (2022). Optimizing the health management information system in Uttar Pradesh, India: Implementation insights and key learnings. *Global Health: Science and Practice*, 10(4).

Moloko, S. M., & Ramukumba, M. M. (2022). Healthcare providers' views of factors influencing family planning data quality in Tshwane District, South Africa. *African Journal of Primary Health Care & Family Medicine*, 14(1), 1-10.

Moses, T. D. S., Kaunda, Z. K., & Ezeron, W. B. (2019). *Analysing, interpreting, and communicating routine family planning data in South Sudan*.

Mucee, E.M., Kaburi, W., Odhiambo-Otieno, P.G., & Kinyamu, R.K. (2016). Routine health management information use in the public health sector in Tharaka Nithi County, Kenya. *Imperial Journal of Interdisciplinary Research*, *2*(3), 660-672.

Musa, S. M., Haruna, U. A., Manirambona, E., Eshun, G., Ahmad, D. M., Dada, D. A., Gololo, A. A., Musa, S. S., Abdulkadir, A. K., & Lucero-Prisno III, D. E. (2023). Paucity of health data in Africa: An obstacle to digital health implementation and evidence-based practice. *Public Health Reviews*, 44. https://doi.org/10.3389/phrs.2023.1605821

Nansikombi, H. T., Kwesiga, B., Aceng, F. L., Ario, A. R., Bulage, L., & Arinaitwe, E. S. (2023). Timeliness and completeness of weekly surveillance data reporting on epidemic prone diseases in Uganda, 2020–2021. *BMC Public Health*, *23*(1), 647.

Nguefack-Tsague, G., Tamfon, B. B., Ngnie-Teta, I., Ngoufack, M. N., Keugoung, B., Bataliack, S. M., & Bilounga Ndongo, C. (2020). Factors associated with the performance of routine health information system in Yaoundé-Cameroon: A cross-sectional survey. *BMC Medical Informatics and Decision Making*, 20, 1-8.

Ngugi, P. N., Gesicho, M. B., Babic, A., & Were, M. C. (2020). Assessment of HIV data reporting performance by facilities during EMR systems implementations in Kenya. *In 18th annual International Conference on Informatics, Management, and Technology in Healthcare (ICIMTH 2020*), held virtually in Athens, Greece, from 3–5 July 2020 (Vol. 272, pp. 167-170).

Orr, K. (1998). Data quality and systems theory. *Communications of the ACM*, *41*(2), 66-71.

Roomaney, R. A., Pillay-van Wyk, V., Awotiwon, O. F., Nicol, E., Joubert, J. D., Bradshaw, D., & Hanmer, L. A. (2017). Availability and quality of routine morbidity data: review of studies in South Africa. *Journal of the American Medical Informatics Association*, *24*(e1), e194-e206.

Rumisha, S. F., Lyimo, E. P., Mremi, I. R., Tungu, P. K., Mwingira, V. S., Mbata, D., ... & Mboera, L. E. (2020). Data quality of the routine health management information system at the primary healthcare facility and district levels in Tanzania. *BMC medical informatics and decision making*, *20*, 1-22.

Rumunu, J., Wamala, J. F., Sakaya, R., Konga, S. B., Igale, A. L., Adut, A. A., ... & Talisuna, A. O. (2022). Evaluation of integrated disease surveillance and response (IDSR) and early warning and response network (EWARN) in South Sudan 2021. *The Pan African Medical Journal*, *42*(Suppl 1).

Sako, S., Gilano, G., Chisha, Y., Shewangizaw, M., & Fikadu, T. (2022). Routine health information utilization and associated factors among health professionals working in public health facilities of the South Region, Ethiopia. *Ethiopian Journal of Health Sciences*, 32(2), 433-444. https://www.ajol.info/index.php/ejhs/article/view/223351

Sanjel, K., Sharma, S. L., Gurung, S., Oli, M. B., Singh, S., & Pokhrel, T. P. (2024). Quality of routine health facility data for monitoring maternal, newborn and child health indicators: A desk review of DHIS2 data in Lumbini Province, Nepal. *Plos One*, *19*(4), e0298101.

Shama, A. T., Roba, H. S., Abaerei, A. A., Gebremeskel, T. G., & Baraki, N. (2021). 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*, 21(1), 1-12.

Shamba, D., Day, L. T., Zaman, S. B., Sunny, A. K., Tarimo, M. N., Peven, K., ... & Lawn, J. E. (2021). Barriers and enablers to routine register data collection for newborns and mothers: EN-BIRTH multi-country validation study. *BMC Pregnancy and Childbirth*, *21*, 1-14.

Shaw, H.L. (2018). Intervening variables. In: Vonk, J., Shackelford, T. (eds). *Encyclopedia of Animal Cognition and Behavior*. Springer: Cham. https://doi.org/10.1007/978-3-319-47829-6_1047-1

Shiferaw, A. M., Zegeye, D. T., Assefa, S., & Yenit, M. K. (2017). 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*, 17(1), 1-9.

Stedman, C., & Vaughan, J. (2022). Data quality. Data management. https://www.techtarget.com/searchdatamanagement/definition/data-quality

Suman, A. (2014). Developing conceptual frameworks: Evolution and architecture. *From Knowledge Abstraction to Management*, 87–108. https://doi.org/10.1533/9781780633695.87

Tavakol, M., & Wetzel, A. (2020). Factor Analysis: a means for theory and instrument development in support of construct validity. *International Journal of Medical Education*, *11*, 245.

Teklegiorgis, K., Tadesse, K., Terefe, W., & Mirutse, G. (2016). Level of data quality from Health Management Information Systems in a resource limited setting and its associated factors, Eastern Ethiopia. *South African Journal of Information Management*, 18(1), 1-8.

Tilahun, H., Abate, B., Belay, H., Gebeyehu, A., Ahmed, M., Simanesew, A., ... & Wondarad, Y. (2022). Drivers and barriers to improved data quality and data-use practices: an interpretative qualitative study in Addis Ababa, Ethiopia. *Global Health: Science and Practice*, *10*(Supplement 1).

Tulu, G., Demie, T. G., & Tessema, T. T. (2021). Barriers and associated factors to the use of routine health information for decision-making among managers working at public hospitals in North Shewa Zone of Oromia Regional State, Ethiopia: A mixed-method study. *Journal of Healthcare Leadership*, 157-167.

Vaismoradi, M., Turunen, H., & Bondas, T. (2013, March 11). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, 15(3), 398–405. https://doi.org/10.1111/nhs.12048

Wandera, S. O., Kwagala, B., Nankinga, O., Ndugga, P., Kabagenyi, A., Adamou, B., & Kachero, B. (2019). Facilitators, best practices and barriers to integrating family planning data in Uganda's health management information system. *BMC Health Services Research*, 19, 1-13.

What is Systems Theory? - Social work theories. (2023). CORP-MSW1 (OMSWP). https://www.onlinemswprograms.com/social-work/theories/systems-theory-social-work/

WHO (2022). *Assessment of Sudan's health information system 2020*. https://applications.emro.who.int/docs/9789290229681-eng.pdf?ua=1

World Health Organization. (2008). *Framework and standards for country health information systems*. World Health Organization.

World Health Organization. (2010). *Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies*. World Health Organization.

World Health Organization. (2017). *Data quality review: A toolkit for facility data quality assessment.* Geneva: World Health Organization.

Wude, H., Woldie, M., Melese, D., Lolaso, T., & Balcha, B. (2020). Utilization of routine health information and associated factors among health workers in Hadiya Zone, Southern Ethiopia. *Plos One*, 15(5), e0233092.

Xie, C. X., Sun, L., Ingram, E., De Simoni, A., Eldridge, S., Pinnock, H., & Relton, C. (2023). Use of routine healthcare data in randomised implementation trials: A methodological mixed-methods systematic review. *Implementation Science*, 18(1). https://doi.org/10.1186/s13012-023-01300-4

Appendix 1: Consent Note & Questionnaire

Consent Note

I am __________ (state name and place of work).

I am doing interviews with department heads and health workers today to learn more about factors influencing the quality of data at the health facility.

You are invited to take part in this interview because the details you provide will help in determining the factors influencing the quality of data in healthcare. During this interview, you will have the chance to share your experiences which will help us identify any gaps and develop improvement strategies.

The interview should **last 10-35minutes.**

No dangers or discomforts are anticipated when you answer the questions.

You can be confident that any information you provide, none of your personal information will be recorded and that it will be kept secret. Your involvement is voluntary.

Do you agree to participate?

NB: ***Only proceed with the interview if the person agrees to be interviewed.***

Part I: - Tool for behavioural and organizational assessment

A self-administered survey for health professionals

Regarding the interviewer: educate participants about the study and obtain their informed consent before beginning the investigation

98	________/________/____________ DD/MM/YYYY
Type of healthcare facility;	
Name of the Health facility	
Unit	
Tel No: (Office)	

COMPLETENESS OF HEALTH FACILITIES REPORTING TO County Health Department

Qn99a How many monthly RHIS reports are supposed to be submitted to the CHD by the health facility?

(SPECIFY THE NUMBER OF REPORTS ACCORDING TO THE FACILITY TYPE)

Health facility type	Number of reports submitted
1. Hospitals	
2. Health centres	
3. Health Units	

Qn99b Are all the monthly RHIS reports in the health facility submitted to the CHD for the following months?

(CHECK THE MONTHLY RHIS REPORTS SUBMITTED BY THE HEALTH FACILITIES DURING THE REVIEW PERIOD)

A. Month 1	B. Month 2	C. Month
1. Yes	1. Yes	1. Yes
2. No	2. No	2. No

REPORT TIMELINESS

Qn100a	1. Is there a deadline for submission of the monthly RHIS report by the health facilities to the CHD?	1. Yes 2. No
Qn100b	2. If yes, what is the Reporting deadline? ______________________________	

Qn100c	Are monthly RHIS reports submitted on time to CHD in the following months? (CHECK THE SUBMISSION DATES OF THE AGGREGATE RHIS REPORTS FOR THE THREE REVIEW MONTHS)

A. Month 1	B. Month 2	C. Mo
1. Yes	1. Yes	1. Ye
2. No	2. No	2. No

Behavioural factors

Qn101. Gender

1. Female

2. Male

Qn102. The respondent's age in years

a) less than 25 years

b) 26-40 years

c) 41-56 years

d) Above 56 years

Qn103. Position of person interviewed

1. Health facility in charge
2. Head of department
3. HMIS Focal Person (Data Clerk
4. Others ………………………………..

Qn104. Level of education attained

1. None
2. Certificate
3. Diploma

4. Bachelor Degree

5. Master Degree

a. Other…………………..

Qn105. Respondent education's field of study

a. CHW/MCHW

b. Nurse

c. Midwife

d. Clinical medicines

e. Other……………………………….

Qn106. Years of experience at work?

a. 1-6 years

b. 7-12 years

c. Above 12 years

Qn107. Are there any rewards or motivations for the HMIS process?

a. Yes

b. No

Qn108. If yes, what type of rewards?

a. Cash

b. Training

c. Appreciation

d. Others……………………………………

Qn109. Have you taken part in the collection or aggregation of data from the health facility's registration form or tally sheet?

a. Yes

b. No

Organizational factors

Qn110. In the past 6 months, have you ever gotten training on HMIS operations?

a. Yes

b. No

Qn111. Have there been review meetings conducted to discuss data quality performance?

a. Yes

b. No

Qn112. And if so, how often?

a. monthly

b. Quarterly

c. Annually

Qn113. Have you received data quality supervision from the county health department or the state in the last three months?

a. Yes

b. No

Qn114. If yes, for Q16, how frequent is the supervision?

a. Monthly

b. Quarterly

c. Annually

Qn115. Did the supervisor evaluate the accuracy of the data?

a. Yes

b. No

Qn116. If so, did the supervisor evaluate the data quality using a checklist?

a. Yes

b. No

Qn117. Did you get regular feedback from the County health department or state on data quality?

a. Yes

b. No

Qn118. If yes, how often?

a. Monthly

b. Quarterly

c. Annually

Technical factors

Qn119. Availability of standard data collection tools?

Yes

No

Qn120. How often are you supplied with data collection tools?

a. Depending on need

b. Monthly

c. Quarterly

d. Others (specify)……………………………………….

Qn121. Do you have a set of indicators that are standardized and defined?

a. Yes

b. No

Qn122. Do you believe that the report and registration forms are user-friendly and simple to understand?

a. Yes

b. No.

Qn123. Do you have qualified human resources who can fill out formats?

a. Yes

b. No

Appendix 2: Key Informant Interview Guide

1. Describe the data collection processes in this health facility (documentation to reporting)

2. Are there practices put in place to guarantee the health facility's data quality?

3. From your knowledge and experience in the health facility, what are the main reasons that promote data quality (good or poor) in the health facility?

4. What recommendations would you make to enhance the quality of the data at this facility?

Appendix 3: Letter of Approval from the University

OFFICE OF THE DEAN, GRADUATE SCHOOL

January, 13 2023

Lubang Denson SHS/MPH/4093-2/2021

Proposal Title: Factors Influencing Data Quality in Routine Health Information System at Health Facilities in Maridi County- South Sudan.

Following your full proposal presentation on 8th September 2022, and subsequent review of your revised proposal, Graduate School has approved your work for submission for ethical review before the commencement of fieldwork. Ethical approval is a mandatory requirement for all research process.

You are advised to seek ethical approval from your country Institutional Research Board (IRB) and share the copy of approval letter with Graduate School.

You are required to update Graduate School of your progress after every three months by submitting progress reports using the forms attached.

Dr Dancan Irungu
Dean Graduate School

CC: HOD Community Health

Amref International University
Langata Road

P.O. Box 27691 - 00506
Nairobi, Kenya

Tel: +254 20 699 3000
Fax: +254 20 600 9518

enquiry@amref.ac.ke
www.amref.ac.ke

Appendix 4: Letter of approval from the Research Ethics Review Board

REPUBLIC OF SOUTH SUDAN

Ministry of Health, Research Ethics Review Board (MOH-RERB), Juba.

MOH/RERB/ Protocol: 16/28/01/2023 **Date:** 22nd /02/2023

Project Title: "Factors Influencing Data Quality in Routine Health Information System at Health Facilities in Maridi County"

MOH/RERB/Approval/16/22/02/2023

Principal Investigator(s): Lubang Denson, MPH Student, Amref University,

Notice of Research Approval

This served to inform you that the research intervention described in the submitted protocol have been reviewed with comments, ethical opinions and suggestions to be harmonized by the Principal Investigator and thereof, the Research Ethics Review Board(MOH-RERB) of the Ministry of Health has determined that according to the National Guidelines September 2019, for research involving humans in the Republic of South Sudan, the activity highlighted therein, meets the requirement and criteria for approval for the implementation of the research activity and exempted from its MOH-RERB oversight.

The MOH-RERB requires you to comply with all institutional guidelines, rules, and regulations and with the tenets of the code. MOH-RERB reserves the right to conduct compliance visit to your research site without previous notification. This approval valid till 30th April 2023.

MINISTRY OF HEALTH, RESEARCH ETHICS REVIEW BOARD MOH-RERB 22 FEB 2023 JUBA - SOUTH SUDAN

Amanya Jacob, MPH-SM, PhD Cand. NU,

D/Director Research, MOH-Juba

Deputy Chairperson, Research Ethics and Review Board, Ministry of Health-Juba, Republic of South Sudan (RERB-MOH)

Cc: U/S MOH-Juba,

Cc: D/G PPB & R, MOH-Juba

Cc: D/G SMOH WES, & Maridi County Health Department/HFs

Tel: +211920536030 Email: ministryofhealthrerb@gmail.com

Appendix 5: Missing values

Variable	Number	Missing Count	Missing Percentage
Name of health facility	106	0	0%
Type of health facility	106	0	0%
Qn99a	12	0	0%
Qn99b	12	0	0%
Qn100a	12	0	0%
Qn100b	12	0	0%
Qn100c	12	0	0%
Qn101	106	0	0%
Qn102	106	0	0%
Qn103	106	0	0%
Qn104	106	0	0%
Qn105	106	0	0%
Qn106	106	0	0%
Qn107	106	0	0%
Qn108	74	32	43%
Qn109	106	0	0%
Qn110	106	0	0%
Qn111	106	0	0%
Qn112	81	25	31%
Qn113	106	0	0%
Qn114	85	21	25%
Qn115	106	0	0%
Qn116	86	20	23%
Qn117	106	0	0%
Qn118	62	44	71%
Qn119	106	0	0%
Qn120	106	0	0%
Qn121	106	0	0%
Qn122	106	0	0%
Qn123	106	0	0%

Appendix 6: Codebook for factors influencing data quality

Theme	Subtheme	Questions	Codes
Data Quality	a. Data quality practices in the facility	1. Describe the data collection processes in this health facility 2. Are there practices put in place to guarantee the health facility's data quality?	a.1. Data collected from the facility register and entered into monthly reports a.2.1 Yes, such as data verification by the person in charge a.2.2 crosschecking the reports before submission a.2.3 Prepare reports jointly with the team to avoid errors a.2.4 Documenting all information about the patients
	1. Behavioral factors influencing data quality	3.From your knowledge and experience in the health facility, what are the main reasons that promote data quality (good or poor) in the health facility? 4.What recommendations would you make to enhance the quality of the data at this facility?	1.3.1 lack of motivation 1.3.2 negative attitude towards work 1.3.3.work overload 1.3.4 lack of cooperation among staff 1.4.1 recruiting more staff to the facilities 1.4.2 motivating staff 1.4.3 staff should be cooperative in the workplace
	2. Organizational factors influencing data quality		2.3.1 Inadequate human resources 2.3.2 poor supportive supervision 2.3.3.No training of staff on HMIS 2.3.4 lack of performance feedback to facilities 2.4.1 recruiting enough human resources 2.4.2 Frequent supportive supervision to staff 2.4.3 provide performance feedback to the facility 2.4.4 Trained staff on HMIS

	3. Technical factors influencing data quality		3.3.1 lack of data collection and reporting tools 3.3.2 lack of trained staff on HMIS 3.4.1 provision of data collection and reporting tools 3.4.2 Ensured trained staff are available to perform HMIS tasks

amref INTERNATIONAL UNIVERSITY
INSPIRING LASTING CHANGE

Appendix 7: Similarity Report

FACTORS INFLUENCING DATA QUALITY IN ROUTINE health information System-Final.

ORIGINALITY REPORT

12%	12%	4%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	etd.uwc.ac.za Internet Source	1%
2	hdl.handle.net Internet Source	1%
3	erepository.uonbi.ac.ke Internet Source	1%

Appendix 8: Evidence of Publication

6/26/24, 10:13 AM Gmail - SAJIM 1856: Manuscript Accepted for Publication, Sent to Editing

Gmail **Lubang Denson Justin <dlubangjust@gmail.com>**

SAJIM 1856: Manuscript Accepted for Publication, Sent to Editing

aosis@sajim.co.za <aosis@sajim.co.za> Mon, Jun 24, 2024 at 2:50 PM
Reply-To: Ms Margo Van Blerk <10ts.srsupport@sajim.co.za>
To: dlubangjust@gmail.com, maggwande@gmail.com, tdmssokiri@gmail.com

****

Ref. No.: 1856
Manuscript title: Factors influencing data quality in routine health information systems at health facilities in Maridi county, South Sudan
Journal: South African Journal of Information Management

****

Dear Lubang Morris, Nyongesa Margaret, Tobijo Sokiri

We are pleased to confirm your manuscript's acceptance for publication on 24-Jun-24.

We can also confirm that the Submission and Review Department released your manuscript to our Finalisation Department to commence the various editing processes to secure online publication within the next 90 days (if not sooner).

Kindly note:
1. If you need to make contact with AOSIS Publishing during the finalisation stage of your manuscript, kindly contact us per email or phone.
2. The finalisation procedure works as follows: (a) The first stage is the language editing that is returned to the corresponding Author for review. This will be the final opportunity for the corresponding Author to make text changes to the manuscript. (b) At a later stage, the editorial staff will send the corresponding author one set of galley proofs, at which time the Author will have two working days to mark any typographical errors.
3. Manuscript tracking is available on the submitting authors' journal profile. The submitting Author could visit their home page frequently to assess the stage of the manuscript.

Thank you for your continued patience and support, and we hope you have joined our online community by signing up to our RSS alerts and Twitter page.

Kind regards,
Ms Van Blerk
AOSIS
Supervisor
Submissions and Review Unit
Scholarly Publishing Department
AOSIS Publishing, Empowering Africa through access to knowledge

---

South African Journal of Information Management | https://sajim.co.za
| ISSN: 2078-1865 (PRINT) | ISSN: 1560-683X (ONLINE)