Evaluating DICOM Compliance for Medical Images in Public Health Facilities in Zambia

By

Elijah Chileshe

Supervisors:

Mr. Claytone Sikasote
Dr. Lighton Phiri

A report submitted to the University of Zambia in partial fulfilment of the requirements of the degree of Postgraduate Diploma in Computer Science
THE UNIVERSITY OF ZAMBIA
LUSAKA
2023

Declaration

I, Elijah Chileshe, do hereby declare that the work in this report has not been previously submitted in candidature for any degree. This report is the result of my own work and investigations, except where otherwise stated. Other sources are acknowledged by given explicit references. A complete list of references is given.

Signature………………… Date ……………………

Copyright

Certificate of Approval

This report of Elijah Chileshe has been approved as fulfilling the requirements or partial fulfilment of the requirements for the award of Postgraduate Diploma in Computer Science by the University of Zambia;

Examiner 1	Signature	Date
……………………	……………………	……………………
Examiner 2	Signature	Date
……………………	……………………	……………………
Examiner 3	Signature	Date
……………………	……………………	……………………
Chairperson/ Board of Examiners	Signature	Date
……………………	……………………	……………………
Supervisor	Signature	Date
……………………	……………………	……………………

Abstract

This paper focuses on a detailed examination of the compliance of Medical Image Metadata with the Digital Imaging and Communications in Medicine (DICOM) standard, specifically within the context of public health facilities in Zambia. The analysis delves into the intricacies of ensuring DICOM compliance for medical images, addressing challenges related to the organization of diverse medical image data, navigating ethical considerations surrounding data authorization and anonymization, and implementing detailed procedures for metadata extraction. The primary objective is to provide a clear understanding of the methodologies involved in large-scale analysis of medical image metadata, emphasizing the compliance of DICOM standards within the public health sector in Zambia. This includes the meticulous collection of image sets from various sources, addressing ethical concerns associated with patient privacy and extracting crucial metadata from DICOM files.

Key Terms: *DICOM, Medical Images, Medical Imaging, Metadata, Metadata Analysis*

Acknowledgements

I would like to express my heartfelt gratitude to several individuals who played instrumental roles in the completion of this research report.

First and foremost, I extend my sincere appreciation to my supervisors, Mr. Claytone Sikasote and Dr. Lighton Phiri, their unwavering support, guidance, and encouragement pushed me to go the extra mile, fostering an environment of continuous improvement and self-betterment.

I am profoundly thankful to my parents, Mr. and Mrs Chileshe, for their enduring support throughout this academic journey. Their encouragement, understanding, and belief in my capabilities have been a constant source of motivation.

I would also like to express my heartfelt thanks to my siblings, Peter, Mapalo, and Grace, for their support, encouragement, and uplifting moments. A special note of appreciation goes to Victorious Daka for the invaluable encouragement provided. Your support has meant a lot, contributing to a positive journey.

To everyone mentioned, your contributions have made a significant impact, and I am truly grateful for the support and inspiration you have provided on this academic journey.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

CPU	Central Processing Unit
CRISP-DM	Cross-Industry Standard Process for Data Mining
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
DX	Digital Radiography
GPU	Graphics Processing Unit
GUI	Graphical User Interface
HPC	High-Performance Computing
LMUTH’s	Levy Mwanawasa University Teaching Hospital‘s
ML	Machine Learning
MRI	Magnetic Resonance Imaging
UNZA	University of Zambia
UTH’s	University Teaching Hospital's

CHAPTER ONE

INTRODUCTION

1.1 Introduction

Medical imaging plays a pivotal role in unveiling the complications of internal structures beneath the skin and bones, serving as an indispensable tool for diagnosing abnormalities and treating various diseases [1]. In the realm of healthcare facilities, a number of cutting-edge technologies, including but not limited to X-rays, Magnetic Resonance Imaging (MRI), Computed Tomography scans (CT), and Ultrasound, have ushered in an era of unprecedented access to detailed medical image data. This information holds the key to unravelling patterns that can significantly enhance patient care.

However, the sheer magnitude of this data presents a formidable challenge. The enormity of medical image datasets necessitates innovative approaches to analysis for meaningful insights. In response to this pressing need, this research endeavours to discover practical and scalable methods for the analysis of medical images. The focal point of our investigation lies in the effective processing and comprehension of these images, leveraging advanced techniques such as machine learning (ML).

By harnessing the power of machine learning, we aim to extract valuable insights from the wealth of medical image data. The application of sophisticated algorithms and computational methodologies holds the promise of not only uncovering hidden patterns but also providing a deeper understanding of the complexities within these images. The outcomes of our study have the potential to make significant contributions to the field of medical research, empowering healthcare personnel with the knowledge and tools to make informed decisions for the well-being of their patients.

1.1.1 DICOM Standard

The DICOM (Digital Imaging and Communications in Medicine) standard is a widely adopted framework in the healthcare industry for encoding, exchanging, and managing medical images and associated data [2]. It ensures interoperability and consistency by providing rules and protocols for the acquisition, storage, transmission, and display of images. DICOM standardises the format of image data and metadata, including patient demographics, study information, imaging modalities, acquisition parameters, and clinical annotations. This standardised approach enables seamless integration and analysis of medical images across different systems and facilitates comprehensive understanding and interpretation of the images within a larger healthcare ecosystem.

1.1.2 DICOM Hierarchy

The DICOM (Digital Imaging and Communications in Medicine) standard establishes a hierarchical structure for organizing medical image data, which has implications for the extraction of metadata. The DICOM hierarchy comprises various levels, including patient, study, series, and instance [3]. At the top level, the patient level, information such as patient demographics and unique identifiers is stored. The study level contains data related to a specific medical study, including imaging modalities and study-specific details. Within a study, multiple series can exist, representing different sets of images acquired during the study. Finally, each series consists of individual instances, which are the actual images captured by the imaging equipment. Understanding the DICOM hierarchy is crucial for accurately extracting metadata, as different levels contain distinct sets of information. Metadata extraction processes need to navigate this hierarchy to retrieve relevant data from each level, ensuring comprehensive and accurate analysis of medical image metadata. Figure 1 displays the four levels of DICOM hierarchy information.

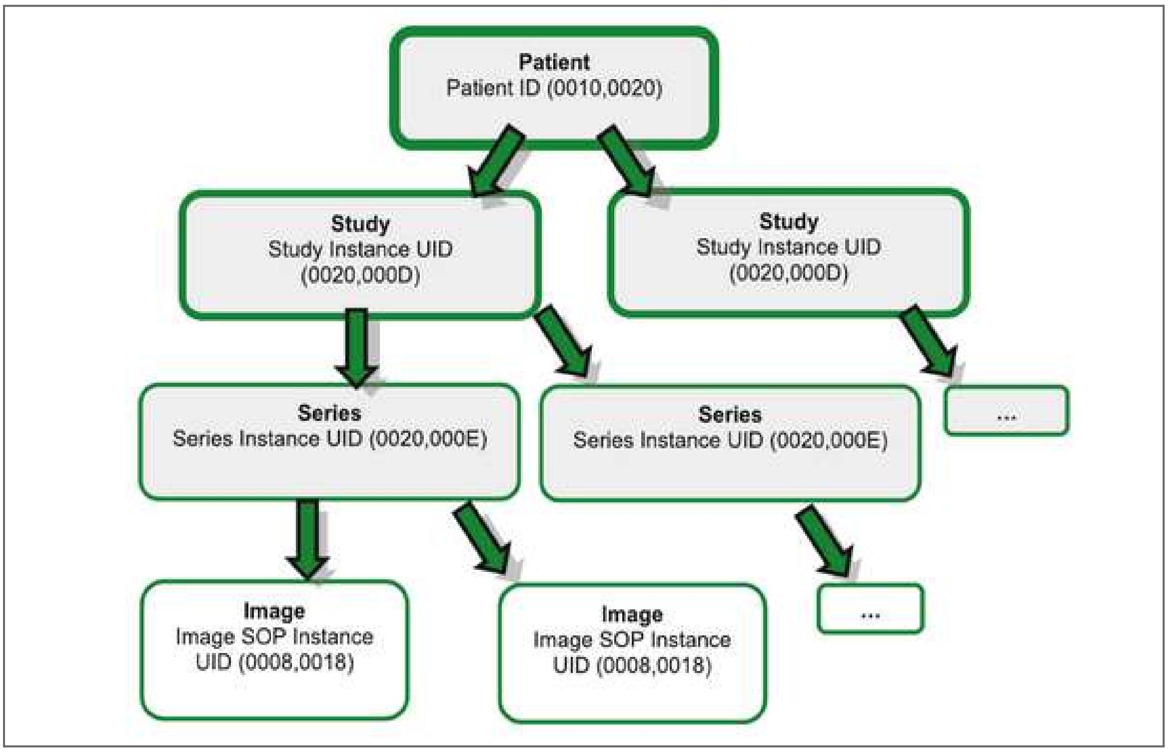

Figure 1. Four levels of DICOM information hierarchy [3]

1.1.3 DICOM Metadata

DICOM metadata encompasses the descriptive information associated with medical images. This includes a wide range of data elements that provide essential context and details about the images [4]. DICOM metadata may include information such as patient demographics (e.g., name, age, sex), imaging modality used (for example, X-ray, CT scan, MRI), imaging acquisition parameters (e.g., exposure settings, image resolution), study information (e.g., study description, study date), and clinical annotations (e.g., radiologist's observations or diagnoses). Extracting DICOM metadata involves parsing the structured data within DICOM files and retrieving specific data elements of interest. This extracted metadata provides valuable insights for large-scale analysis, enabling researchers to study patterns, trends, and associations within medical image data. Understanding the structure and content of DICOM metadata is essential for conducting meaningful analysis and utilizing the full potential of medical image datasets. Table 1 displays a table with a partial representation of DICOM metadata elements.

Table 1. Partial Representation DICOM Metadata Elements

Attribute Name	Tag	Type	Attribute Description
Modality	0008,0060	1	Device that produced the Instances in this Series
Study Description	0008,1030	3	Classification of the Study performed.
Patient Name	0010,0010	2	Patient's full name.
Patient ID	0010,0020	2	Primary identifier for the Patient.
Series Instance UID	0020,000E	1	Unique identifier of the Series.

1.2 Background of Study

This study examines the systematic analysis of Medical Image Metadata using the DICOM standard. It delves into the organization of diverse medical image data, explores ethical considerations related to data privacy and investigates steps for processing and extracting metadata. Amid the backdrop of Zambia's radiologist scarcity, where a few professionals serve a significantly large population, the study emphasizes addressing challenges concerning compliance of medical image metadata to the DICOM standard which is crucial for interpretation of medical images. A central focus are the challenges faced by healthcare practitioners when interpreting images with missing metadata, leading to difficulties in accurate diagnosis and decision-making. Overall, this study provides comprehensive insights into methodologies for large-scale analysis of medical image metadata, aiming to refine imaging practices and inform clinical decisions.

1.3 Problem Statement

The core problem in this study pertains to the compliance of medical image metadata to the DICOM standard, particularly within the context of Zambia's constrained radiologist resources. Despite the advancements in medical imaging technologies, there exists a pressing challenge in ensuring that medical image metadata aligns accurately, completely, and consistently with the DICOM standard.

The scarcity of radiologists in Zambia heightens this issue, emphasizing the need for a meticulous examination of the compliance of medical image metadata to the DICOM standard. The study specifically aims to address the consequences of inadequate compliance, focusing on instances where medical images lack essential metadata. Such deficiencies not only hinder the accuracy of diagnoses but also impede informed decision-making by healthcare practitioners.

By concentrating on the meticulous examination of compliance with the DICOM standard, the research seeks to uncover practical solutions and methodologies for refining the handling and interpretation of medical image metadata. This focused exploration is essential for ensuring that medical imaging practices align with established standards, ultimately contributing to improved patient care and clinical decision-making in Zambia's healthcare landscape.

1.4 Objectives of the Study

The primary objective of this project is to look into practical strategies for conducting a comprehensive analysis of medical images on a large scale. The core focus involves devising efficient techniques to handle and carefully evaluate these images with a high degree of precision and correctness. By harnessing the capabilities of advanced computational methods, our overarching aim is to derive meaningful and valuable insights from these medical images. These insights have the potential to play a pivotal role in enhancing diagnostic capabilities and facilitating well-informed decision-making within the medical field.

1.4.1 General Objectives

The main objective of the study is to investigate the compliance of Medical Image Metadata within medical image practices.

1.4.2 Specific Objectives

1. To analyze the compliance of medical image metadata.
2. To identify the challenges associated with analysis of medical image metadata.
3. To propose recommendations and strategies for enhancing the accuracy and efficiency of medical image metadata

1.4.3 Research Questions

1. How complete is medical image DICOM metadata maintained within the current medical imaging processes?
2. What are the challenges associated with the analysis of medical image metadata?

3. What practical strategies and solutions can be suggested to ensure better compliance with Medical Image metadata, thereby improving the accuracy and efficiency of medical imaging practices.

1.5 Chapter Summary

Chapter One provides an introduction to the research, highlighting the crucial role of medical imaging in healthcare and the challenges posed by the vast amounts of medical image data. The study aims to address these challenges through the systematic analysis of medical image metadata, with a specific focus on compliance with the DICOM standard. The DICOM hierarchy and metadata elements are introduced, emphasizing their significance in ensuring interoperability and comprehensive understanding of medical images. The background of the study contextualizes it within Zambia's radiologist scarcity, underscoring the need for meticulous examination of metadata compliance. The problem statement identifies the core issue of inadequate compliance and its consequences in the context of limited radiologist resources. The objectives of the study encompass a comprehensive analysis of medical image metadata compliance, identification of challenges, and proposing strategies for enhancement.

CHAPTER TWO

RELATED WORK

2.1 Introduction

The use of medical imaging technologies has revolutionised the diagnosis and treatment of various medical conditions. However, the accurate and consistent alignment of medical image metadata with the DICOM standard remains a pressing challenge. The DICOM standard is a widely accepted standard for medical image metadata, which ensures interoperability and consistency across different imaging modalities and healthcare providers. Despite the importance of the DICOM standard, compliance with this standard remains a challenge. This literature review aims to provide insights into the analysis of DICOM metadata.

2.1.1 Automated DICOM metadata and volumetric anatomical information extraction for radiation dosimetry

The paper Automated DICOM metadata and volumetric anatomical information extraction for radiation dosimetry presents a novel approach to extract metadata and volumetric anatomical information from DICOM files for individualized dosimetry [5]. The authors highlight the importance of accurate and complete metadata extraction for patient-specific dosimetry calculations using MC simulation techniques. They propose a user-friendly GUI developed in MATLAB environment that can automatically extract metadata from every slice image of a DICOM file in a single query and interactively specify the regions of interest (ROI) without explicit access to the radiology information system. The developed GUI is fast, easy, and constitutes a useful tool for individualized dosimetry.

The software tool was developed on MATLAB environment [6]. The structure of the GUI has been designed to extract the information in a user-friendly and time-saving functionality. The GUI can automatically extract metadata from every slice image of a DICOM file in a single query and interactively specify the regions of interest (ROI) without explicit access to the radiology information system. The volumetric maps are formed by interactively specifying the ROIs and by assigning a specific value in every ROI. The result is stored in DICOM format, for data and trend analysis.

The developed GUI is easy, fast, and constitutes a very useful tool for individualized dosimetry. The authors successfully demonstrated the feasibility of their approach by extracting metadata and volumetric anatomical information from DICOM files for individualized dosimetry. The metadata are automatically formatted and presented to the user as a Microsoft Excel file. The volumetric maps are formed by interactively specifying the

ROIs and by assigning a specific value in every ROI. The result is stored in DICOM format, for data and trend analysis.

The strengths include the paper proposing a novel approach to extract metadata and volumetric anatomical information from DICOM files for individualized dosimetry. The developed GUI is easy, fast, and constitutes a very useful tool for individualized dosimetry. The authors successfully demonstrated the feasibility of their approach by extracting metadata and volumetric anatomical information from DICOM files for individualized dosimetry.

Some weaknesses include the paper not discussing the limitations of their approach in terms of the accuracy and completeness of metadata extraction and the authors did not compare their approach with other existing DICOM extraction metadata tools. The paper does not discuss the potential impact of their approach on the adherence and compliance of medical image metadata to the DICOM standard.

2.1.2 DICOM in digital dentistry

This paper provides a comprehensive literature review of the role of DICOM in digital dentistry. The authors discuss the impact of digital radiography on dental imaging and the importance of adhering to DICOM standards to ensure reliable transmission of information. They also highlight the need for further standardization in areas such as digital photographic displays and surgical workflow issues [7]. The paper provides a detailed overview of the current state of digital dentistry and the role of DICOM in ensuring reliable transmission of information .

The methodology used in this paper involves a literature review of relevant studies and articles on the topic of DICOM in dentistry. The authors used a comprehensive literature review methodology to analyze the current state of digital dentistry and the role of DICOM in ensuring reliable transmission of information. They discussed the impact of digital radiography on dental imaging and the importance of adhering to DICOM standards to ensure reliable transmission of information [8]. The paper provides clear recommendations for further standardization in areas such as digital photographic displays and surgical workflow issues. The authors also discussed the challenges of ensuring accurate, complete, and consistent alignment of medical image metadata with the DICOM standard and provided potential solutions to address these challenges.

The results of this paper highlight the importance of adhering to DICOM standards in digital dentistry. The authors discuss the challenges of ensuring accurate, complete, and consistent alignment of medical image metadata with the DICOM standard. They also provide recommendations for further standardization in areas such as digital photographic displays and surgical workflow issues. The paper emphasizes the need for continued efforts to ensure that medical image metadata aligns accurately, completely, and consistently with the DICOM standard .

Strengths of this paper include a comprehensive overview of the role of DICOM in digital dentistry. The authors provide a detailed analysis of relevant studies and articles on the topic. The paper also provides clear recommendations for further standardization in areas such as digital photographic displays and surgical workflow issues.

Weaknesses of this paper include a limited discussion of specific challenges in ensuring adherence and compliance of medical image metadata to the DICOM standard. The paper does not provide a detailed analysis of the potential solutions to address these challenges. Additionally, the paper has a limited discussion on the impact of non-compliance with DICOM standards on patient care and outcomes.

2.1.3 DICOM for quantitative imaging biomarker development

This paper delves into the use of the DICOM standard and open-source tools for encoding research data in the development of quantitative imaging biomarkers. It demonstrates that the DICOM standard can effectively represent various analysis results and their relationships, with consistent annotation of data objects using widely accepted codes for semantic communication. Building on the groundwork laid by previous research groups and projects like 3D Slicer and DCMTK, the paper underscores the crucial role of standardization for ensuring accuracy and reproducibility. [9] emphasizes PET image acquisition and quantitative data analysis standards, and [10] introduces the Unified Medical Language System (UMLS) for integrating biomedical terminology. They advocate for open data in science, reducing research waste, and address challenges related to inaccessible research data, proposing solutions to enhance accessibility and transparency.

The authors used the real-life research scenario of HNC PET/CT quantitative image analysis to demonstrate the capabilities of the DICOM standard [10], [11]. Most of the methods used for QI analysis that produced the data presented in this paper are accompanied by FOSS tools developed as part of the QIICR project. The authors focus on the use of DICOM to enable structured, standardized, and interoperable communication of the annotated analysis results produced by those tools. They chose DICOM as the common standard, and demonstrated that it is interoperable with the variety of tools readily available to the researcher, as well as commercial clinical imaging and analysis systems. The paper presents a detailed investigation of the development and application of the DICOM standard and supporting FOSS tools to encode research data for quantitative imaging biomarker development. The authors also demonstrate that the DICOM standard is capable of representing various types of analysis results and their interrelationships, and that the resulting data objects are annotated in a standard manner, utilizing consistent and widely used codes for communicating semantics.

Some weaknesses of the paper focuses on the use of DICOM to enable structured, standardized, and interoperable communication of the annotated analysis results produced by FOSS tools developed as part of the QIICR project, and does not discuss these analysis methods in detail, or validate the tools implementing those analysis methods. The paper also does not address the challenges of ensuring that medical image metadata aligns accurately, completely, and consistently with the DICOM standard, which is the focus of the user's paper.

2.1.4 Outcomes Metadata from Indexing Initiatives of Medical Analysis Imaging DICOM Repositories

This is a paper that reports on a secondary analysis of eight research studies that indexed DICOM metadata from the PACS of different healthcare facilities. The paper discusses the advantages and challenges associated with indexing and managing large volumes of DICOM metadata, and provides valuable insights into the potential of DICOM metadata for secondary analysis of large volumes of medical imaging data [12].

The methodology used in the paper involved the installation of Dicoogle on a personal computer or on virtual machines belonging to the healthcare facilities. The indexed metadata volume ranged between 250 GB and 34.2 TB, which had an impact on indexing time, ranging from 16.5h to 86 days. The paper provides several references to research studies that reported the indexing of DICOM metadata from PACS.

The results of the paper indicate that indexing DICOM metadata from PACS can provide several advantages, such as improving the accuracy and completeness of medical image metadata, facilitating data sharing and interoperability, and enabling secondary analysis of large volumes of medical imaging data. However, the paper also identifies several challenges associated with indexing and managing DICOM metadata, such as the need for standardization and quality control, the complexity of DICOM metadata, and the need for efficient storage and retrieval systems. Overall, the paper provides a comprehensive analysis of the advantages and challenges associated with indexing and managing DICOM metadata, which can inform the development of strategies to ensure compliance with the DICOM standard

The strengths of the paper include the paper providing a comprehensive analysis of the advantages and challenges associated with indexing and managing DICOM metadata, which can inform the development of strategies to ensure compliance with the DICOM standard.

The paper also highlights the potential of DICOM metadata for secondary analysis of large volumes of medical imaging data, which can facilitate research and clinical decision-making.

Weaknesses of the paper include the paper not specifically addressing the issue of compliance of medical image metadata to the DICOM standard. The paper does not provide a detailed methodology for the secondary analysis of the research studies, which may limit the generalizability of the results. The paper does not discuss the potential impact of indexing and managing DICOM metadata on patient privacy and data security, which is an important consideration in the context of medical imaging.

2.1.5 Using DICOM Metadata for Radiological Image Series Categorization: a Feasibility Study on Large Clinical Brain MRI Datasets

The paper Using DICOM Metadata for Radiological Image Series Categorization: a Feasibility Study on Large Clinical Brain MRI Datasets explores the feasibility of using DICOM [13] metadata to automate the identification of brain MRI sequences. The authors provide background information on previous works related to series categorization and DICOM, highlighting the challenges related to series identification in the clinical workflow. They also discuss the two approaches to organizing and categorizing medical imaging acquisitions, which are using metadata or image contents. The authors propose a methodology that includes a fast and efficient series labeling process for training, a DICOM attributes selection strategy and a machine learning step.

The methodology proposed by the authors involves using DICOM metadata to automate the identification of brain MRI sequences. The authors used two large brain MRI datasets from different institutions and continents to test their approach. They also developed a fast and efficient series labeling process for training, a DICOM attributes selection strategy, and a machine learning step that includes the processing of DICOM attributes to input features and the training of a classifier.

The authors tested their approach on two large brain MRI datasets and achieved high efficiency and accuracy in identifying relevant series for further image-related algorithms. They also found that the proposed methodology can be used to identify relevant series for different types of brain MRI studies, including T1-weighted, T2-weighted, and FLAIR sequences .

Strengths of this paper include the fact that it relies on the DICOM standard, which is a widely accepted international data standard for medical imaging. This ensures that the metadata used in the methodology aligns accurately, completely, and consistently with the DICOM standard. Additionally, it is efficient and accurate in identifying relevant series for further image-related algorithms, which can be useful in ensuring that medical image metadata is compliant with the DICOM standard.

However, a potential weakness is that it only focuses on brain MRI sequences, limiting its generalizability to other types of medical imaging acquisitions. Additionally, the authors did not discuss the potential limitations of using DICOM metadata for series identification, such as the variability and unreliability of the series description attribute, which may impact compliance with the DICOM standard.

2.1.6 Role of DICOM in Artificial Intelligence for skin disease

The paper Role of DICOM in Artificial Intelligence for skin disease discusses the potential benefits of using DICOM in artificial intelligence (AI) for skin disease diagnosis and treatment. The authors highlight the importance of standardized image formats and metadata, metadata-based image retrieval, and de-identification protocols in the efficient curation of multi-institutional datasets for machine learning training, testing, and validation. The authors also discuss the potential for AI algorithms to improve diagnostic accuracy and reduce healthcare costs [14].

The methodology used in this study is a literature review of existing research on the topic. The results suggest that DICOM can play a crucial role in the development and implementation of AI for skin disease diagnosis and treatment. The authors also highlight the importance of addressing technological, ethical, regulatory, medicolegal, and workforce barriers before DICOM and AI can be used effectively in dermatology.

One of the key benefits of using DICOM in AI for skin disease diagnosis and treatment is the potential for improved diagnostic accuracy. AI algorithms can analyze large datasets of medical images and identify patterns that may not be visible to the human eye. This can lead to earlier and more accurate diagnoses, which can improve patient outcomes and reduce healthcare costs. Another benefit of using DICOM in AI is the potential for more efficient curation of multi-institutional datasets. The authors note that DICOM mechanisms such as standardized image formats and metadata, metadata-based image retrieval, and de-identification protocols can improve the efficiency of medical imaging workflows and facilitate the sharing of medical images across institutions.

However, there are also several challenges and limitations of the paper AI for skin disease diagnosis and treatment. One of the key challenges is ensuring compliance with the DICOM standard. There is a significant lack of compliance with the DICOM standard in medical imaging, which can lead to inaccurate and inconsistent metadata. Another challenge is the potential for bias in AI algorithms. AI algorithms may be biased if they are trained on datasets that are not representative of the population being diagnosed. Additionally, there are ethical and regulatory challenges associated with the use of AI in healthcare, including concerns about patient privacy and the need for regulatory oversight.

2.1.7 A DICOM Framework for Machine Learning and Processing Pipelines Against Real-time Radiology Images

A DICOM Framework for Machine Learning and Processing Pipelines Against Real-time Radiology Images presents a framework called Niffler that enables efficient retrieval of images from hospitals' PACS for use in research clusters. The authors highlight the challenge of real-time execution of machine learning (ML) pipelines on radiology images due to limited computing resources in clinical environments, and the need for efficient data transfer capabilities when running them in research clusters. The paper presents three use cases of Niffler, including executing an IVC segmentation pipeline, training a deep learning model for COVID-19 detection, and extracting radiomics features for predicting treatment response in glioblastoma patients [15].

The methodology used in the paper involves the development of Niffler, an open-source framework that leverages DICOM metadata to enable efficient retrieval of images from hospitals' PACS for use in research clusters. The authors describe the architecture of Niffler, which consists of three main components: the PACS connector, the data transfer module, and the processing module. The PACS connector retrieves DICOM images and metadata from the hospital's PACS, while the data transfer module ensures fast and secure transfer of data to the research cluster. The processing module sorts the received DICOM images and facilitates the execution of ML pipelines .

The results observed in the paper demonstrate the effectiveness of Niffler in enabling real-time execution of ML pipelines on radiology images. The authors report that Niffler was able to execute an IVC segmentation pipeline in less than 10 seconds, which is significantly faster than the traditional approach of manually transferring images to a research cluster. The authors also report that Niffler was able to train a deep learning model for COVID-19 detection using a dataset of 1,000 chest X-ray images in less than 10 minutes. Finally, the authors report that Niffler was able to extract radiomics features for predicting treatment response in glioblastoma patients using a dataset of 50 MRI images in less than 5 minutes .

In terms of strengths, the paper presents a novel framework that addresses the challenge of real-time execution of ML pipelines on radiology images in clinical environments. The authors demonstrate the effectiveness of Niffler in enabling efficient retrieval of images from hospitals' PACS for use in research clusters, and present three use cases that highlight the versatility of the framework.

In terms of weaknesses, the paper focuses primarily on the technical aspects of the framework and does not provide a detailed discussion of the potential ethical implications of using patient data for research purposes. Additionally, the paper does not provide a comparison of Niffler with other existing frameworks for retrieving medical images from hospitals' PACS. Finally, the paper does not provide a detailed discussion of the limitations of the framework and the potential challenges that may arise when using it in real-world settings.

2.1.8 Deep Semi-Supervised Algorithm for Learning Cluster-Oriented Representations of Medical Images

The paper Deep Semi-Supervised Algorithm for Learning Cluster-Oriented Representations of Medical Images Using Partially Observable DICOM Tags and Images proposes a novel approach to automatically extracting large homogeneous datasets of medical images based on detailed criteria. The authors use a deep semi-supervised clustering algorithm that leverages both images and partially observable DICOM tag metadata from a fraction of the available data [16]. The proposed model architecture can generalise well, as demonstrated by evaluating model performance on test data using several evaluation methods. The authors also confirm by visual inspection that it groups visually similar images, even when having only partially observable DICOM metainformation.

The authors trained the models using 30,000 images and tested them using a disjoint test set consisting of 8000 images, gathered retrospectively from the PACS repository of the Clinical Hospital Centre Rijeka in 2017. They compared their method against the standard and deep unsupervised clustering algorithms, as well as the popular semi-supervised algorithms combined with the most commonly used feature descriptors. Their model achieved an NMI score of 0.584 with respect to the anatomic region and an NMI score of 0.793 with respect to the modality. The results suggest that DICOM data can be used to generate pairwise constraints that can help improve medical images clustering, even when using only a small number of constraints .

One strength of the paper Deep Semi-Supervised Algorithm for Learning Cluster-Oriented Representations of Medical Images Using Partially Observable DICOM Tags and Images is that it proposes a novel approach to automatically extracting large homogeneous datasets of medical images based on detailed criteria and/or semantic similarity. This approach can be useful in identifying and addressing issues related to compliance with the DICOM standard. Another strength is that the proposed model architecture can generalize well, which is important for ensuring that the extracted datasets are representative of the larger population of medical images. This can help improve the accuracy, completeness, and consistency of medical image metadata with the DICOM standard.

One weakness is that the authors do not provide a comparison of their results with other studies that have used similar methods. This makes it difficult to assess the generalizability of their approach and its potential impact on improving compliance with the DICOM standard. Another weakness is that the authors do not discuss the limitations of their approach, such as the potential for bias in the selection of training and test datasets. This is important for understanding the potential limitations of the proposed approach in addressing compliance issues related to medical image metadata and the DICOM standard.

2.1.9 QBot: A Tool for Monitoring Compliance with Standard Operation Procedures in Nuclear Medicine

The paper QBot: A Tool for Monitoring Compliance with Standard Operation Procedures in Nuclear Medicine presents an innovative tool called Q-Bot, which uses DICOM metadata to monitor compliance with standard operating procedures (SOPs) in nuclear medicine. The authors highlight the complexity and time-consuming nature of regular and precise inspection of the realization of local nuclear medicine SOPs, especially when large amounts of patient data are obtained from a wide range of different scan procedures on a daily basis [17]. The authors argue that Q-Bot can help address this challenge by automatically monitoring relevant parameters, such as patient ID, patient mass and height, injected activity, and uptake time, in the case of adult PET/CT and gamma camera bone scans. The authors also note that

Q-Bot provides a graphical user interface (GUI) that summarizes outliers in a table format to be investigated by a dedicated technologist.

The authors tested Q-Bot for 11 months at two nuclear medicine departments. They collected information related to the error handling for retrospective analysis of long-term tendencies. The authors retrieved header information from real patient DICOM files to detect errors in the clinical workflow as reflected in the metadata. They introduced Q-Bot in the routine clinical workflow of two nuclear medicine departments and monitored DICOM records for several months. The authors used metadata in the header of the DICOM images, which include information about the patient, the applied radiopharmaceutical and its administration, the scanner acquisition settings, and the reconstructed image-related parameters. The authors note that the metadata regularly includes data elements, and each of them has an identifying code and an actual value.

The authors found that Q-Bot was capable of verifying SOP conformities and automatically inspecting the actual SOP compliance of relevant DICOM parameters. The authors also found that Q-Bot's GUI provided a summary of the outliers in a table format to be investigated by a dedicated technologist. The authors note that Q-Bot was introduced in the routine clinical workflow of two nuclear medicine departments and monitored DICOM records for several months. The authors present the most relevant clinical use of Q-Bot that may provide valuable information to other nuclear medicine centers as well.

Some strengths of the paper is that Q-Bot is an innovative tool that uses DICOM metadata to monitor compliance with standard operation procedures in nuclear medicine, which can help ensure that medical image metadata aligns accurately, completely, and consistently with the DICOM standard. Q-Bot is capable of automatically monitoring relevant parameters, such as patient ID, patient mass and height, injected activity, and uptake time, in the case of adult PET/CT and gamma camera bone scans, which can help ensure that medical image metadata aligns accurately, completely, and consistently with the DICOM standard.

Some weaknesses are that the study was conducted at only two nuclear medicine departments, which may limit the generalizability of the findings to other medical imaging modalities and departments. The study did not compare Q-Bot to other tools or methods for monitoring compliance with standard operation procedures in medical imaging, which may limit the ability to determine the most effective approach for ensuring compliance with the DICOM standard.

2.1.10 Self-Supervised Pretraining with DICOM metadata in Ultrasound Imaging

The paper "Self-Supervised Pretraining with DICOM metadata in Ultrasound Imaging" presents a robust approach to leveraging DICOM metadata for self-supervised pretraining in ultrasound image analysis. This methodology addresses the challenge of effectively utilizing the abundant information encoded in DICOM format, aligning with the broader goal of conducting comprehensive analysis of medical images on a large scale. By incorporating DICOM metadata as weak labels, the paper demonstrates a practical strategy for improving representation learning and downstream tasks, thereby enhancing the precision and correctness of medical image analysis [18]. This approach aligns with the overarching aim of deriving meaningful insights from medical images, which can significantly contribute to enhancing diagnostic capabilities and informed decision-making in the medical field. However, the paper’s focus is specifically on ultrasound imaging and may not directly address the broader spectrum of medical image analysis.

A strength of the paper is that it introduces an innovative approach by using DICOM metadata as weak labels for self-supervised pretraining in ultrasound imaging. This innovative methodology demonstrates a practical strategy for improving representation learning, which is in line with the need for advanced techniques to derive meaningful insights from medical images, as outlined in your paper.

The paper's approach has the potential to enhance the accuracy of ultrasound image analysis through improved representation learning. This potential impact aligns with the overarching aim of deriving meaningful insights from medical images, potentially contributing to enhanced diagnostic capabilities and informed decision-making in the medical field, which is a key focus of your paper.

A weakness of the paper includes the paper’s exclusive focus on ultrasound imaging may limit the generalizability of the findings to other medical imaging modalities. This limitation could potentially restrict the applicability of the approach to a broader range of medical image analysis, which may not fully address the need for comprehensive analysis of medical images on a large scale

2.2 Chapter Summary

This chapter offers an extensive overview of the challenges associated with aligning medical image metadata with the DICOM standard. It delves into various applications such as radiation dosimetry, digital dentistry, and quantitative imaging biomarker development. The literature review encompasses different papers, including automated metadata extraction for dosimetry, DICOM's role in digital dentistry, and the use of DICOM in AI for skin disease diagnosis. The chapter also explores a tool for monitoring compliance in nuclear medicine. The strengths and weaknesses of each approach are discussed, highlighting the importance of adherence to the DICOM standard across diverse medical imaging contexts.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

The research methodology outlines the planned procedures for data collection, analysis, and interpretation of the study. It also provides a structured approach to address the research problem and achieve the defined objectives.

3.2 Research Approach

To address the primary objective of this project, a mixed methods approach was employed, incorporating both qualitative and quantitative methods to comprehensively analyze medical images on a large scale.

3.2.1 Qualitative Methods

Qualitative methods were utilized to gather valuable insights from medical image operators regarding the metadata elements intended to be entered during image acquisition. This involved conducting interviews with medical image operators to understand their perspectives on the essential metadata elements during image acquisition.

The qualitative phase aimed to uncover the aspects of metadata entry, providing a qualitative foundation for understanding the operators' perspectives and practices.

3.2.2 Qualitative Methods

The quantitative component of the research involved a careful analysis to assess the level of DICOM compliance within the collected medical images. This encompassed:

1. Metadata Extraction: Employing advanced computational methods for extracting DICOM metadata from the medical images.
2. DICOM Compliance Assessment: Applying standardized criteria and guidelines to evaluate the compliance of extracted metadata with the DICOM standards.

The quantitative phase aimed to provide a systematic analysis of the DICOM metadata, focusing on compliance to the DICOM standard.

By combining both qualitative and quantitative methods, the research approach aimed to offer an understanding of the practical strategies employed by medical image operators and the level of DICOM compliance in large-scale medical image datasets.

3.3 Research Design

In this research, the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework was used as the guiding research design. CRISP-DM is a widely recognized methodology for conducting data mining projects, offering a structured and iterative approach to the entire data analysis process. The CRISP-DM consists of the following phases:

1. **Business Understanding**

In the business understanding phase, the focus is on understanding the research goals. This involves defining the research problem, establishing objectives and gaining insights into the context of the study. For this paper, the business understanding phase corresponds to the identification of key research questions and objectives, aligning with the broader goal of analysis of DICOM Compliance for medical image metadata.

2. **Data Understanding**

The data understanding phase involves exploring and familiarizing oneself with the available data. In the context of this research, it entails obtaining a comprehensive understanding of medical image datasets, including the nature of metadata and variations in imaging modalities. See section 3.3.1.1.

3. **Data Preparation**

The data preparation phase focuses on cleaning, preprocessing and transforming the data to make it suitable for analysis. In the context of this paper, it involves the extraction of metadata from medical images and the preparation of datasets for subsequent analysis. See section 3.3.1.2.

4. **Modelling**

The modeling phase encompasses the application of advanced computational methods to analyze the prepared data. In this research, it involves the quantitative assessment of DICOM compliance within the medical image datasets obtained. Advanced computational techniques are leveraged to help with the analysis. See section 3.3.1.5.

5. **Evaluation**

The evaluation phase assesses the effectiveness and validity of the derived insights. In the context of this paper, it includes evaluating the level of DICOM compliance and the practical implications of missing metadata.

6. **Deployment**

The deployment phase refers to the implementation of models or findings into practical applications. Since this study focuses on analysis and assessment without the deployment of specific tools or applications, the emphasis remains on deriving insights from the medical image metadata.

By adopting the CRISP-DM framework, this research ensures a systematic and structured approach to analyzing medical image metadata on a large scale, providing a foundation for deriving meaningful and valuable insights. The iterative nature of CRISP-DM allows for the refinement of strategies and methodologies throughout the research process, ultimately contributing to the achievement of the research objectives.

3.3.1 Evaluation of Medical Image Metadata Compliance

3.3.1.1 Selection of Data Sources for Comprehensive Medical Image Data Collection and Analysis

The selection of appropriate data sources is a crucial aspect in the collection of medical image data for large-scale analysis. Various sources contribute to the diversity and comprehensiveness of the dataset. These sources may include hospitals, clinics, imaging centers, research institutions, and public databases. Collaboration and partnerships with healthcare providers and institutions are essential to access a wide range of imaging data. It is important to consider factors such as the availability of diverse patient populations, different imaging modalities, and a variety of medical conditions [19]. Furthermore, ensuring the data sources are reliable and representative is vital for the generalizability and validity of the analysis. Obtaining data from multiple sources increases the chances of capturing a comprehensive view of the target population, enabling more accurate and meaningful insights. Careful consideration should also be given to data sharing agreements, data ownership, and compliance with relevant regulations to ensure responsible and ethical use of the collected data. Below are the two institutions where the Medical Images were obtained from:

1. The University Teaching Hospital's
2. Levy Mwanawasa University Teaching Hospital's (LMUTH's)

3.3.1.2 Digitalization of Medical Images

Efforts are being made to digitize medical images as part of the implementation process. This involves converting physical films or analog images into a digital format. The digitization process typically includes scanning or capturing the images using specialized equipment such as digital scanners or medical imaging devices [20]. Once the images are digitized, they can be stored, processed and analyzed more efficiently using computer systems and software. Digitalization allows for easier accessibility, sharing and manipulation of medical images, enabling large-scale analysis of metadata encoded in the DICOM standard. See Fig 2:

Figure 2. Digitalization of Medical Images

3.3.1.3 Leveraging Parallel Processing for Efficient Analysis of Large-Scale Medical Image Metadata

To handle the computational demands of processing large-scale image metadata, parallel processing libraries such as Joblib [21], PySpark [22], and Dask [23] can be employed. These libraries enable the distribution of computational tasks across multiple processors or machines, allowing for faster and more efficient analysis. Joblib, for example, provides tools for parallel computing in Python, allowing tasks to be executed in parallel across multiple cores or even on remote machines. PySpark, on the other hand, is a powerful framework for distributed data processing that utilizes a cluster computing system, making it suitable for processing large datasets in a distributed manner. Dask, similar to PySpark, provides scalable parallel computing capabilities, enabling efficient processing of large-scale image data. This is particularly beneficial when dealing with large volumes of medical image data, where traditional sequential processing may be time-consuming and impractical. Parallel processing allows for the concurrent execution of tasks, effectively reducing the overall processing time and enabling more rapid analysis of the image data.

3.3.1.4 Anonymization of Patient Identifying Information

Leveraging Parallel Processing when anonymizing patient data is important because it helps anonymize the data faster. Anonymization is integral when analyzing patient images because it ensures the removal or encryption of patient-specific details, safeguarding sensitive information. By applying anonymization to patient identifying data, the analysis process maintains privacy compliance standards while allowing for robust and efficient examination of medical images. This protective measure is crucial in handling sensitive healthcare data, aligning with ethical considerations and regulatory requirements.

3.3.1.5 Analysis of Medical Image Metadata

In the research, the Python programming language to conduct image processing tasks on datasets stored across external hard drives. The use of Python facilitated efficient and customizable workflows for handling diverse medical image data. Specifically, for reading and managing DICOM (Digital Imaging and Communications in Medicine) files, a prevalent format in medical imaging, I leveraged the capabilities of the Pydicom library.

Pydicom proved instrumental in the extraction of information from DICOM files, encompassing critical metadata, patient demographics, and pixel data from medical images. Its functionalities empowered me to seamlessly navigate, modify, and analyze the extensive dataset of medical images stored on external hard drives, contributing to the overall success of the research endeavor.

This approach allowed for a flexible and scalable methodology, aligning with the specific requirements of medical image analysis while harnessing the capabilities of Python and Pydicom to process and interpret the intricacies within the datasets

3.3.2 Assessment of Challenges Arising from Analysis of Medical Image Metadata

3.3.2.1 Challenges in Managing and Storing Large Scale Medical Image Data

The large-scale analysis of medical image data presents several challenges that must be overcome to achieve meaningful results. One primary challenge is the management and storage of the immense volume of data generated by medical imaging technologies. As hospitals move towards a filmless, paperless environment, there will be a never-ending demand for digital storage space [24]. Developing efficient storage systems capable of handling the continuous production of data while ensuring data accessibility and integrity is a significant undertaking. This may involve the use of distributed storage solutions, cloud-based storage or data archiving strategies to optimize storage capacity and data retrieval performance.

3.3.2.2 Computational Challenges in Processing Large Scale Medical Image Datasets

Another major challenge is the computational demand associated with processing large-scale image datasets. Processing large datasets requires substantial computational resources and can be computationally expensive and time-consuming. There needs to be access to powerful computing infrastructures equipped with high-performance GPUs [25] or even applying for High Performance Computing (HPC) [26] to accelerate processing time. Implementing techniques such as model parallelism or distributed computing frameworks can help alleviate the computational burden.

3.3.2.3 Ensuring Data Privacy and Security in Medical Image Analysis

Ensuring data privacy and security is an ongoing challenge when working with medical image data. Ethical guidelines encourage respecting privacy, that is, the ability to retain complete control and secrecy about one's personal information [27]. Medical images contain sensitive patient information that must not be disclosed, making it crucial to implement robust data protection measures, adhere to privacy regulations and adopt secure data transfer protocols. Encryption, access controls and anonymization techniques play a vital role in safeguarding patient privacy and maintaining data security throughout the analysis pipeline.

3.3.2.4 Addressing Data Quality and Veracity Challenges in Large-Scale Medical Image Analysis

Data quality and veracity are critical aspects in the analysis of large-scale medical image data. The variability of medical images, stemming from factors like acquisition protocols and imaging modalities, necessitates addressing data quality challenges for accurate analysis. Standardization techniques, preprocessing steps, and robust algorithms mitigate variability and enhance data quality. Additionally, establishing standardized imaging protocols contributes to more consistent and comparable medical images. Veracity, encompassing issues such as inconsistencies, missing data, ambiguities, deception, fraud and duplication is vital in healthcare decision-making, and managing data quality is a fundamental challenge . By addressing data quality and veracity concerns, healthcare professionals can ensure reliable and trustworthy information for improved analysis and decision-making.

3.3.3 Development of Recommendations and Strategies for Improved Accuracy and Efficiency of Medical Image Metadata

This section will focus on developing recommendations and strategies for improving the accuracy and efficiency of medical image metadata. This will involve:

3.3.3.1 Identifying key areas of non-compliance with the DICOM standard

Based on the analysis, specific areas where medical image metadata experiences inconsistencies or deficiencies will be identified. This will involve examining mandatory and recommended DICOM elements, focusing on critical information relevant to diagnosis and clinical decision-making.

3.3.3.2 Developing practical recommendations for metadata improvement

Based on the identified areas of non-compliance, a set of practical and context-specific recommendations will be formulated. These recommendations will address strategies for ensuring accurate, complete, and consistent metadata capture, entry, and management.

3.4 Data Analysis

In this study, information about mandatory DICOM metadata elements was gathered through a data collection method involving interviews. Participants were queried using a tailored questionnaire designed to extract insights specifically related to DICOM metadata. The qualitative nature of the study allowed for a nuanced exploration of the subject matter. Through thematic analysis, major themes were identified, providing a comprehensive understanding of the mandatory DICOM metadata. The resulting data was organized to align with research objectives, enabling meaningful conclusions about the significance and applicability of these elements.

3.5 Ethical Consideration

Ethical considerations are critical in the implementation of large-scale analysis of medical image data. This involves ensuring privacy and confidentiality by obtaining informed consent, de-identifying or anonymizing data, and implementing strict access controls and encryption techniques [28]. Ethical clearance for the research has already been secured from both The University of Zambia Biomedical Research Ethics Committee (Reference Number: 2731-2022) and The National Health Research Authority (Reference Number: NRHA000024/10/05/2022). The study is authorized to use medical image datasets from both the University Teaching Hospitals and Levy Mwanawasa University Teaching Hospital.

3.6 Chapter Summary

The methodology chapter employs a mixed methods approach, combining qualitative insights gathered through interviews with medical image operators and advanced computational methods for quantitative DICOM compliance assessment within medical images. Adhering to the CRISP-DM framework, the structured research design addresses challenges in managing large-scale data and computational demands while prioritizing ethical considerations. Recommendations for improved metadata accuracy stem from identified areas of non-compliance. Additionally, a qualitative inquiry into mandatory DICOM metadata elements through interviews provides nuanced insights aligned with research objectives. Ethical considerations, including approvals from relevant committees, are central to the research process, establishing a comprehensive framework for medical image metadata analysis.

CHAPTER FOUR

EVALUATION

4.1 Evaluation

All the experiments were conducted on a standalone LENOVO ® ThinkPad T480, with an Intel ® Core ™ i5-8350U (CPU @ 1.70GHz), using 16 GB RAM, and running Ubuntu 22.04.6 LTS

4.1.1 Datasets

All datasets that were used for conducting the experiments were obtained from the University Teaching Hospitals and Levy Mwanawasa University Teaching Hospitals.

4.1.2 DICOM files

Only DICOM files were used during the study from both institutions University Teaching Hospitals and Levy Mwanawasa University Teaching Hospitals. See table 2.

Table 2. Total Number of DICOM Files Obtained

Modality	The University Teaching Hospital's	Levy Mwanawasa UTH
X-Rays	4171	2532
CR	1639	0

.

4.1.3 DICOM Fields Analyzed

Some mandatory DICOM fields were used to analyze the DICOM files [29]. Table 3 displays the following fields were utilized:

Table 3. DICOM Metadata Elements Analyzed

Attribute Name	**Tag**	**Type**	**Attribute Description**
Operators Name	0008,1070	3	Name of the Operator
Referring Physicians Name	0008,0090	2	Name of the Patient's referring Physician

Modality	0008,0060	1	Device that produced the Instances in this Series
Patient Name	0010,0010	2	Patient's full name.
Patient ID	0010,0020	2	Primary identifier for the Patient.
Patient Sex	0010,0040	2	Sex of the Patient
Patient Age	0010,1010	3	Age of the Patient
Study Date	0008,0020	2	Date the Study started.
Study ID	0008,1030	3	Classification of the Study performed.
Institution	0008,0080	3	Institution where the equipment that produced the composite instances is located.

4.2 Chapter Summary

This chapter outlines the experimental setup and criteria for assessing the proposed study. Datasets exclusively sourced from University Teaching Hospitals and Levy Mwanawasa University Teaching Hospitals were employed, with a specific focus on DICOM files. The chapter details the quantity of DICOM files obtained, particularly for X-rays and CR. Additionally, the evaluation involved a targeted analysis of essential DICOM fields, emphasizing the careful examination of metadata elements such as Operator's Name, Referring Physician's Name, Modality, Patient Name, Patient ID, Patient Sex, Patient Age, Study Date, Study ID, and Institution.

CHAPTER FIVE

RESULTS AND DISCUSSIONS

5.1. Results of Medical Image Metadata Compliance Assessment

Below are the results for the analysis carried out for the DICOM files from the University Teaching Hospitals and Levy Mwanawasa University Teaching Hospitals:

5.1.1 The University Teaching Hospitals

5.1.1.1 Computed Radiography - 2013

Essential patient identification fields (Patient Name, Patient ID, Study Date, and Study ID) consistently complete at 100%, showcasing robust data capture practices.
Operators Name, Referring Physicians Name, Patients Age and Institution, however, exhibit 0% completeness while Patient Sex has less than 10%, suggesting potential gaps in data entry. See Fig 2 below:

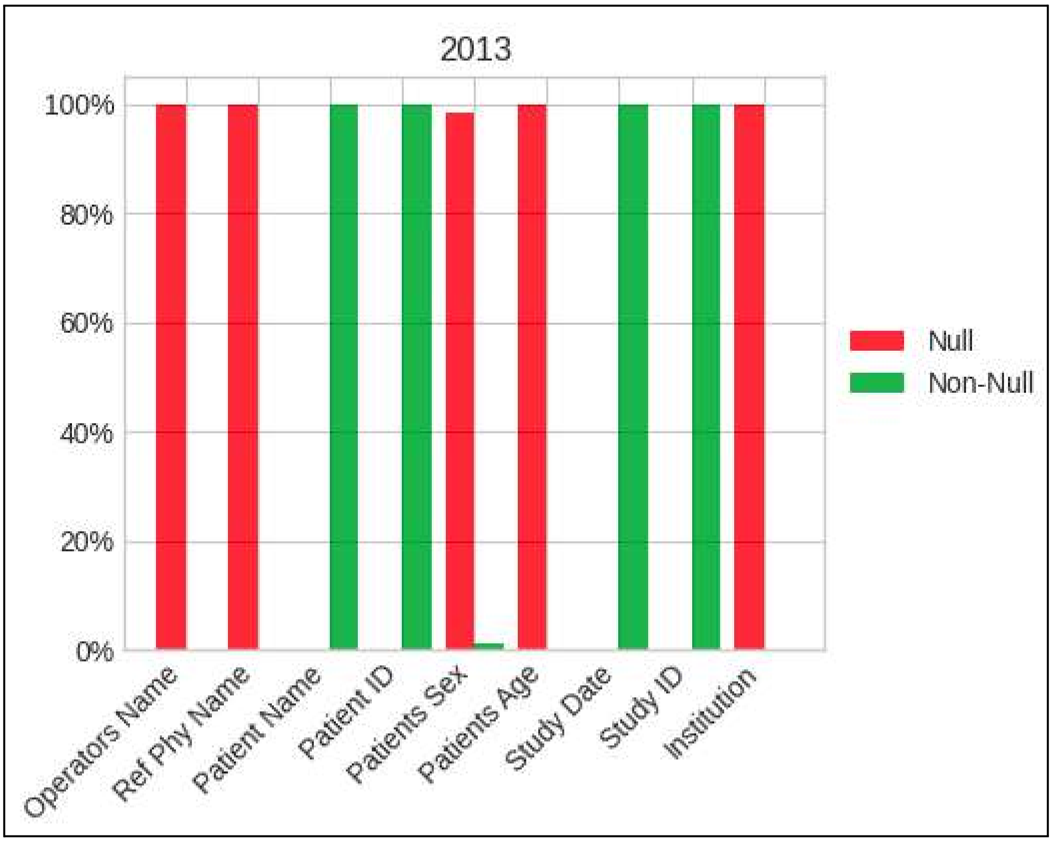

Figure 3. UTH's Computed Radiography - 2013

5.1.1.2 Computed Radiography - 2014

Key fields, including Patient Name, Patient ID, Study Date, and Study ID, maintain full completeness at 100%, reflecting data uniformity.

Operators Name, Referring Physicians Name, Patient Sex, Patient Age and Institution remain at 0%, reinforcing the need for comprehensive data capture practices. See Fig 3 below:

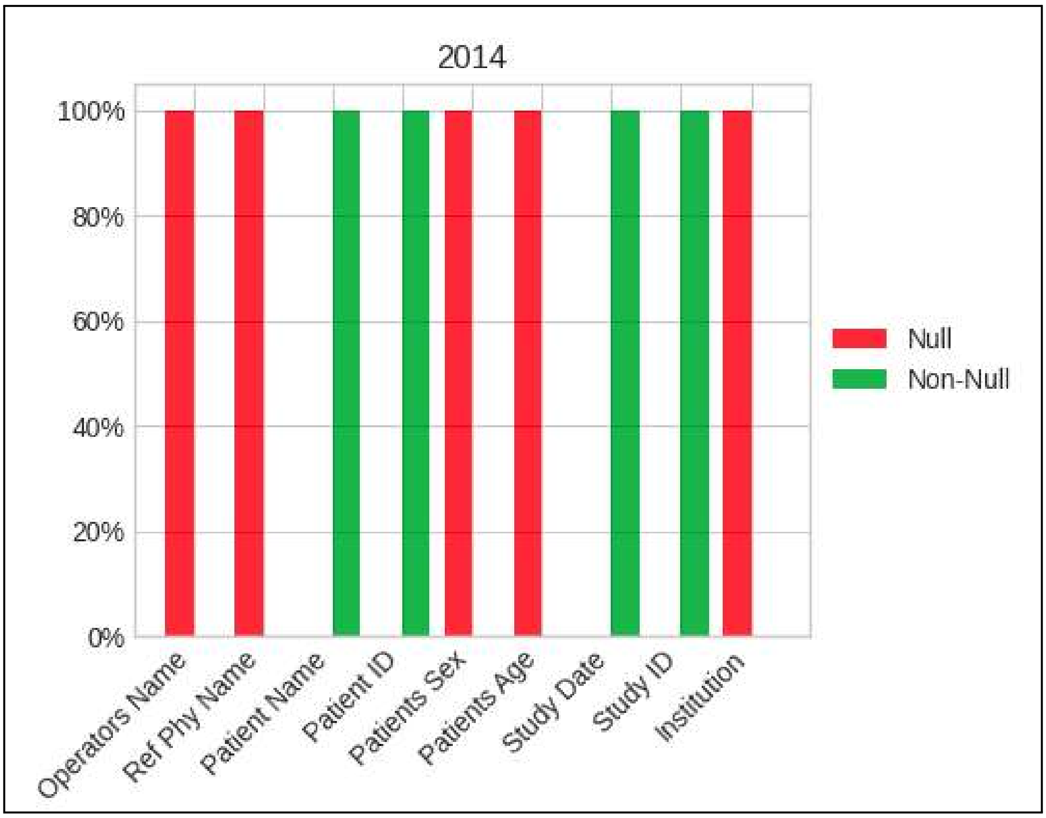

Figure 4. UTH's Computed Radiography - 2014

5.1.1.3 Digital Radiography - 2015

High percentages in critical fields (Patient Name, Patient ID, Study Date) maintain data integrity at 100%, with a notable increase in Institution and Patient Sex completeness at 98%. Operators Name, Referring Physicians Name and Patient Age show 0% completeness, signaling areas for improvement. See Fig 4:

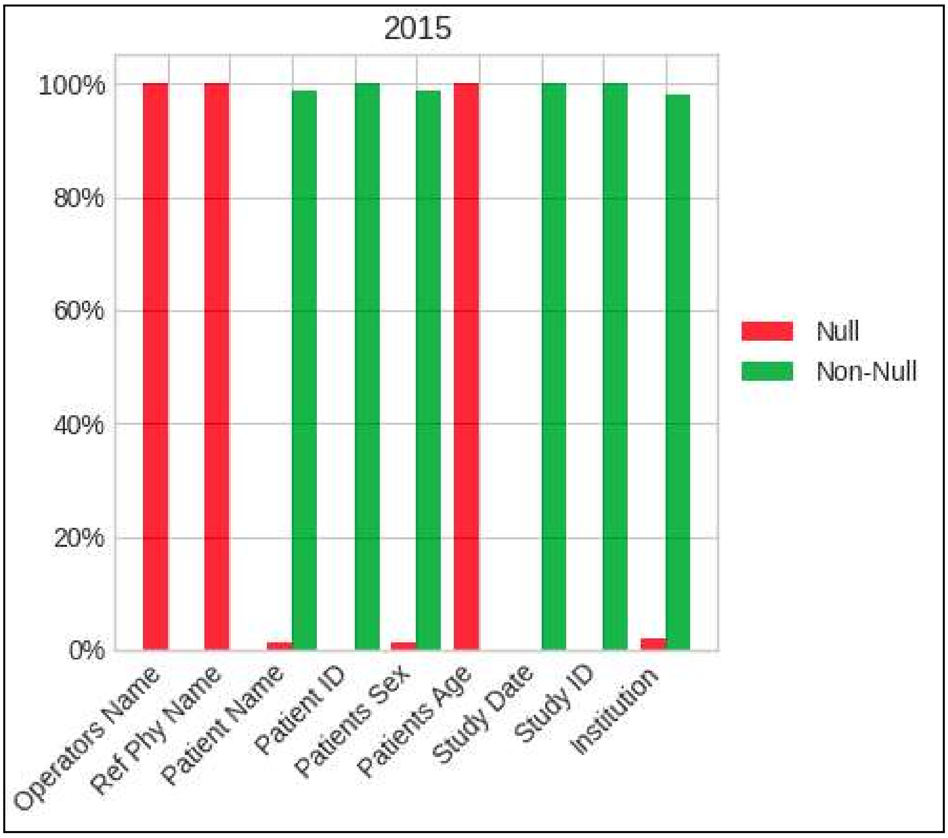

Figure 5. UTH's Digital Radiography - 2015

5.1.1.4 Digital Radiography - 2016

Patient-centric data (Patient Name, Patient ID and Patient Sex) is complete, and although there is a decline in institutional completeness and Patient Sex, it remains considerable.
Study ID and Study date show an increase in metadata completeness at 100%.
Operators Name, Referring Physicians Name, Patient Age and Institution persist at 0%, emphasizing the need for enhanced data entry procedures. See Fig 5 below:

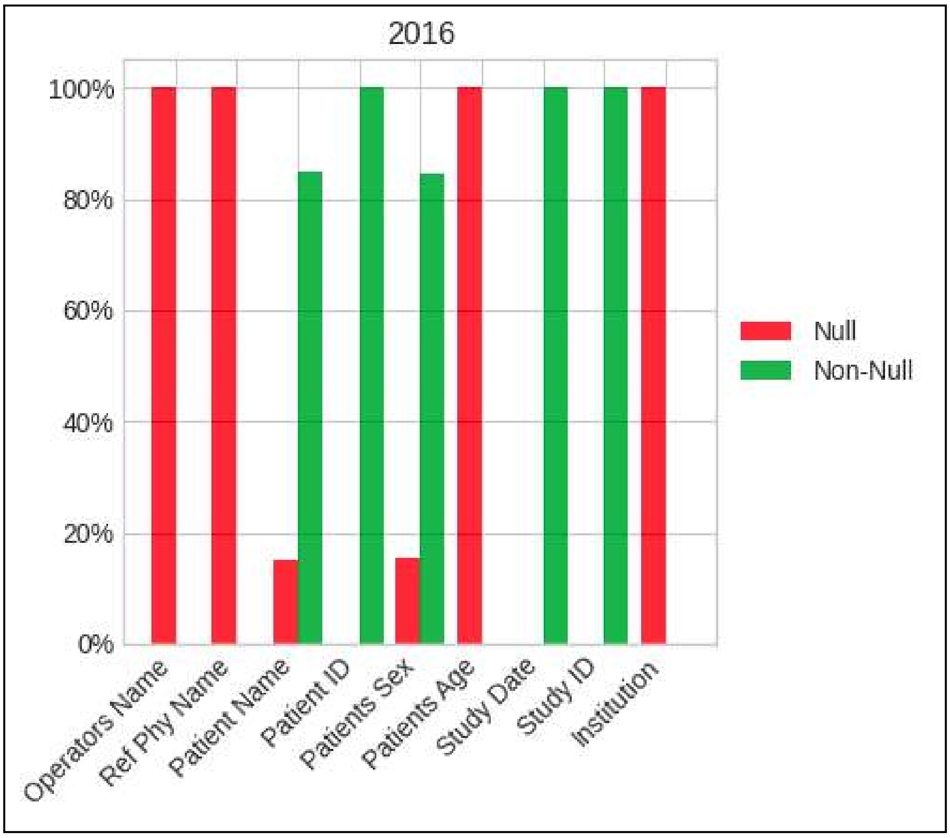

Figure 6. UTH's Digital Radiography - 2016

5.1.1.5 Digital Radiography - 2017

Patient-centric data maintains completeness, while there's a moderate decrease in Patient Sex, indicating a potential shift in data capture practices.

Study ID and Study date show an increase in metadata completeness at 100%.

Operators Name, Referring Physicians Name, Patient Age and Institution remain at 0%, requiring attention to comprehensiveness. See Fig 6 below:

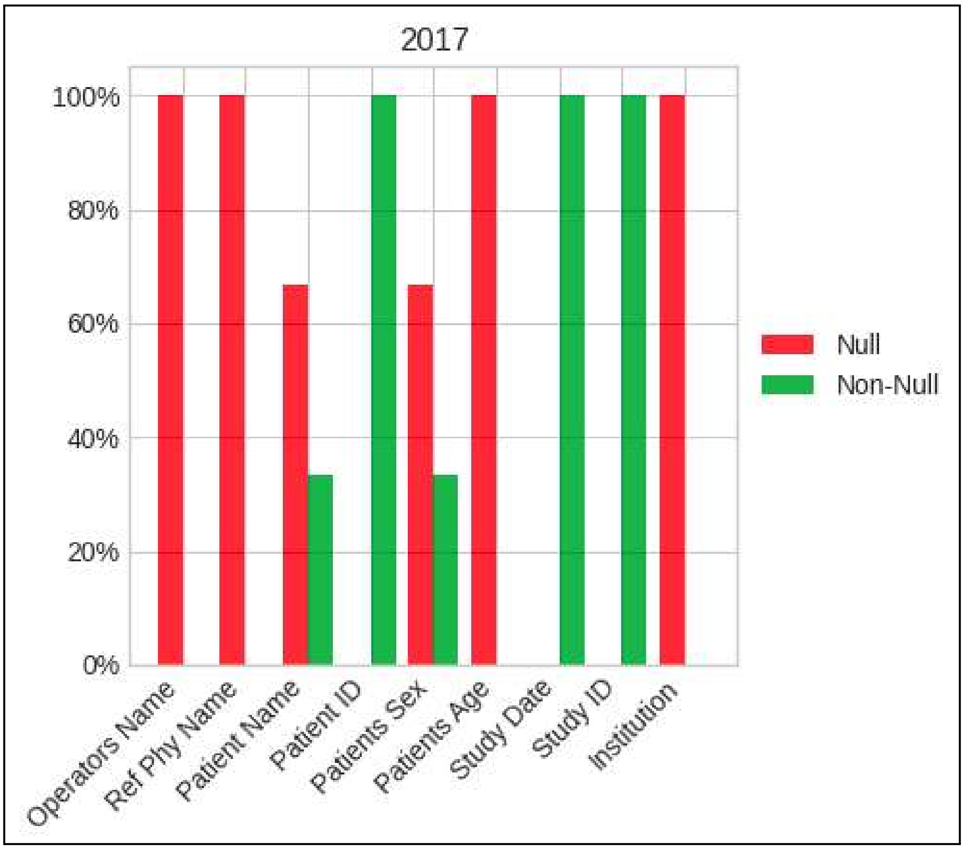

Figure 7. UTH's Digital Radiography - 2017

5.1.1.6 Digital Radiography - 2018

Essential fields (Patient ID, Study Date and Study ID) remain complete at 100%, maintaining a high standard of data capture while Patient Name and Sex are at 98% completeness. Operators Name, Referring Physicians Name, and Institution persist at 0%, pointing towards the need for improved data capture. See Fig 7 below:

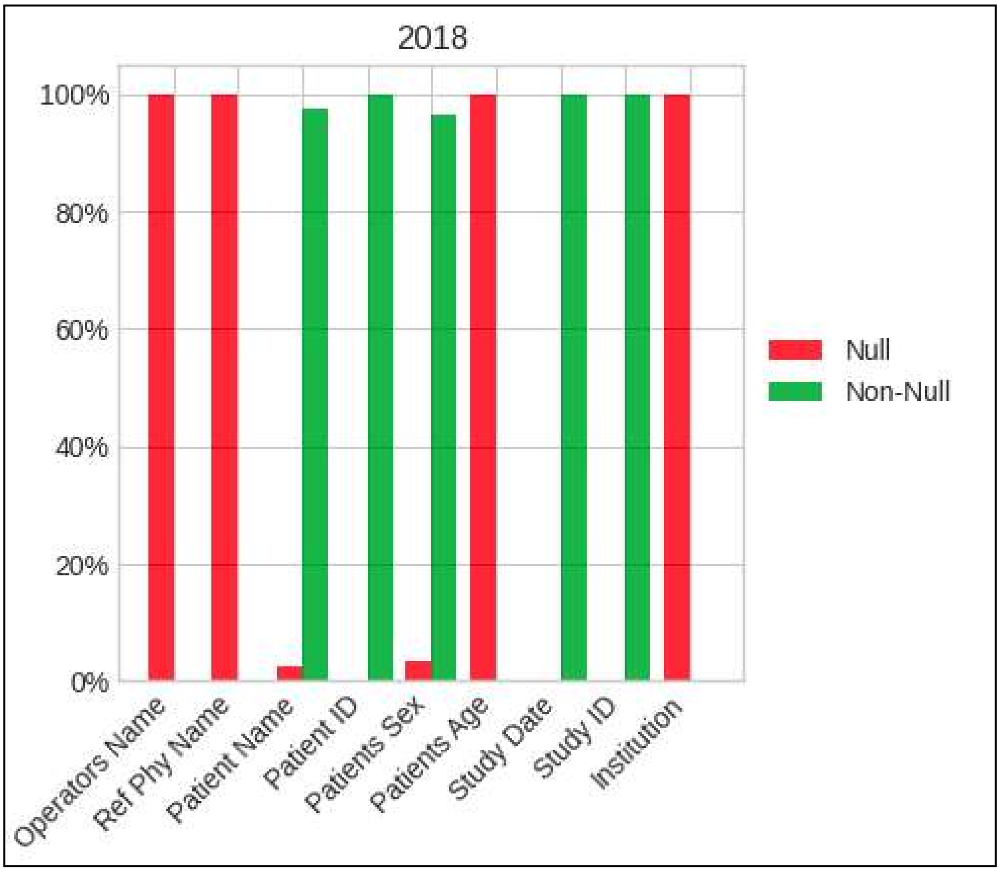

Figure 8. UTH's Digital Radiography - 2018

5.1.2 Levy Mwanawasa University Teaching Hospital

5.1.2.1 Digital Radiography - 2020

Remarkable completeness across all fields (Patient Name, Patient ID, Patient Sex, Patient Age, Study Date, Study ID, and Institution) at 100%.

Referring Physicians Name shows 0% while Operators Name is at 74% completeness, warranting further investigation into data entry processes. See Fig 8 below:

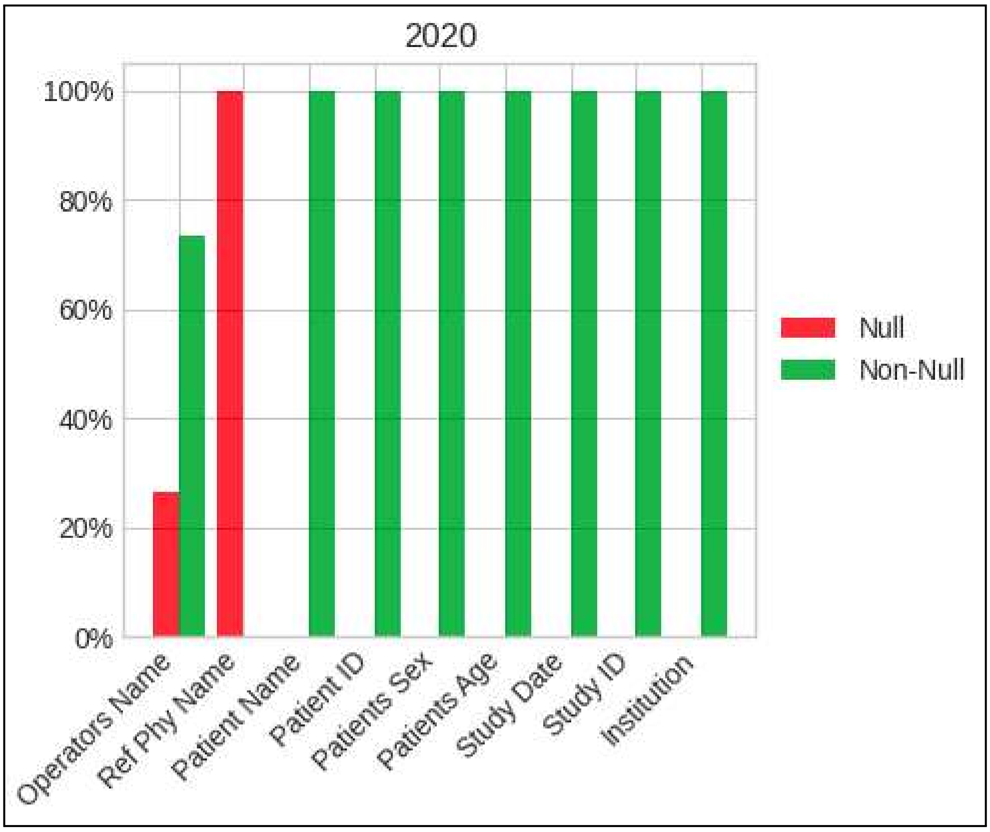

Figure 9. Levy Mwanawasa UTH's Digital Radiography - 2020

5.1.2.2 Digital Radiography - 2021

High completeness in essential fields (Patient Name, Patient ID, Patient Sex, Patient Age, Study Date, Study ID, and Institution) at 100%.

Referring Physicians Name remains at 0% and Operators Name at 50% completeness, necessitating improvements in data extraction or entry procedures. See Fig 9 below:

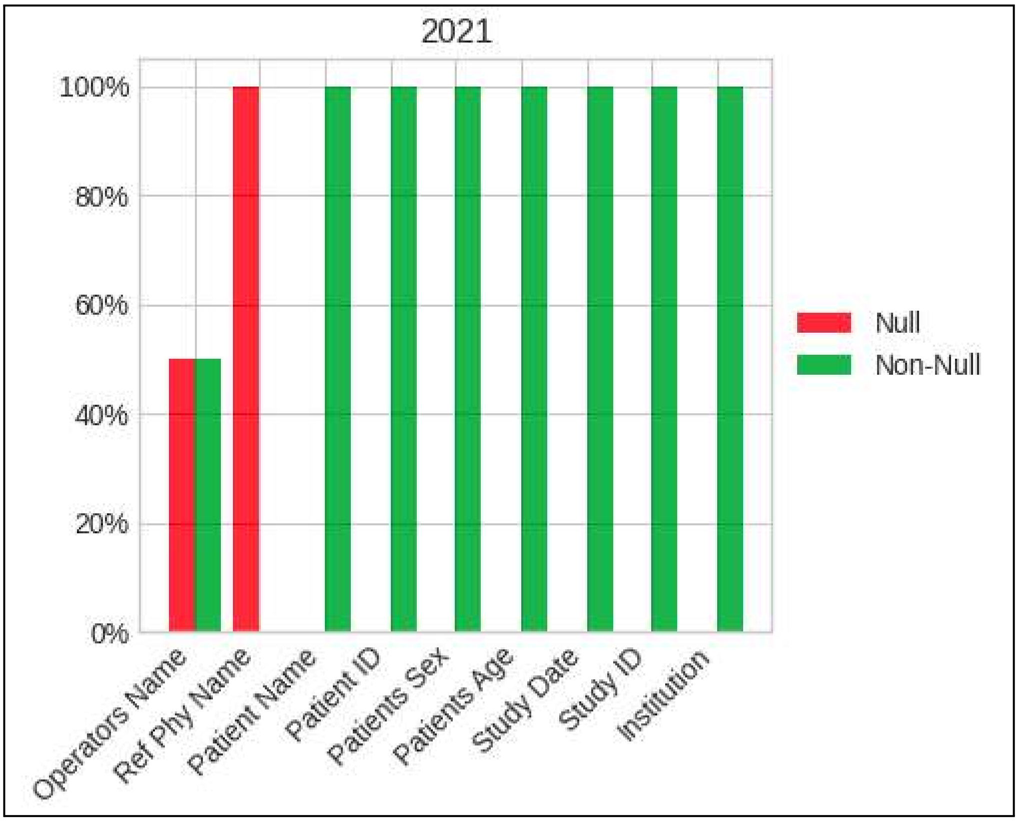

Figure 10. Levy Mwanawasa UTH's Digital Radiography - 2021

5.1.2.3 Digital Radiography - 2015

Exceptional completeness across all fields, indicating a significant improvement in data capture practices Operators Name at 98% while the other fields (Referring Physicians Name, Patient Name, Patient ID, Patient Sex, Patient Age, Study Date, Study ID, and Institution) at 100%.

No 0% completeness in this instance, showcasing comprehensive data capture. See Fig 10 below:

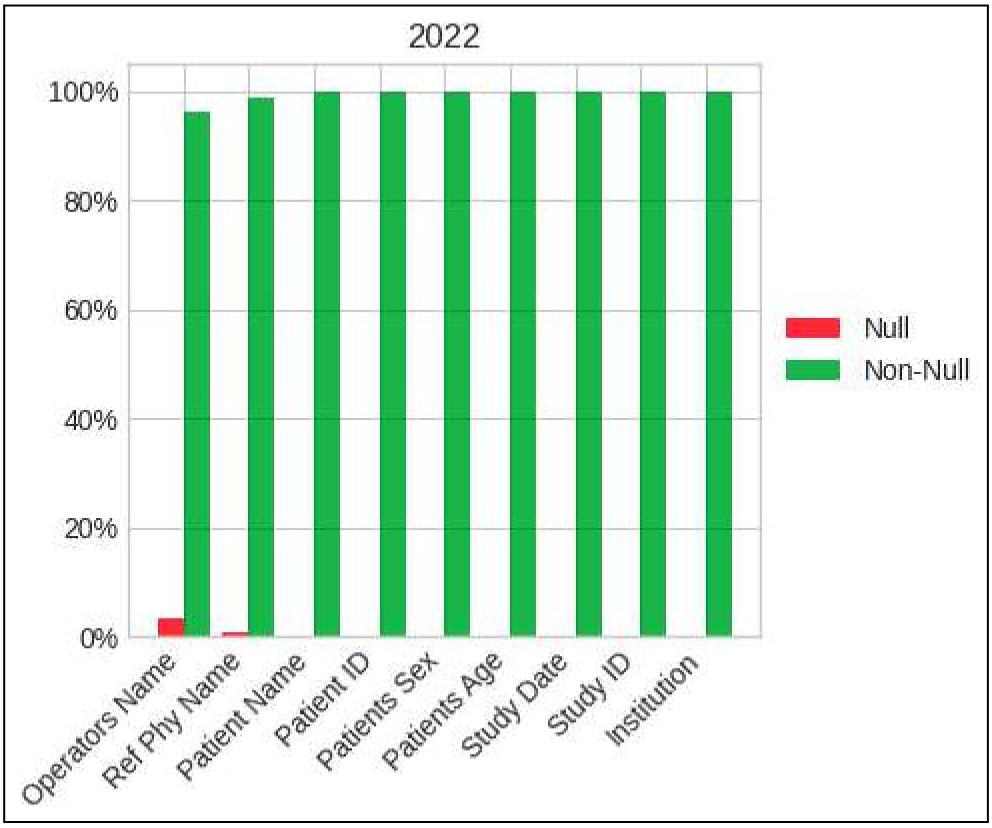

Figure 11. Levy Mwanawasa UTH's Digital Radiography - 2022

5.2 Assessment of Results of Interviews with Radiographers

5.2.1 Additional information entered during an examination using modality machines

The results of the interviews regarding the additional information entered during examinations using modality machines reveal commonalities and variations in participant responses. Participant 1 indicated entering "Examination," "Patient Age," and "Patient Sex" as specific information. On the other hand, Participant 2 mentioned entering "Examination," "Patient Sex," and "Patient Age." The consistent inclusion of "Examination" underscores its importance, while the differing sequence of "Patient Age" and "Patient Sex" suggests some variability in the data entry practices among participants.

5.2.2 Examining Information Sources

The interview results pertaining to the sourcing of information for DICOM metadata fields indicate a consistent approach among participants. Both Participant 1 and Participant 2 reported obtaining information for DICOM metadata fields primarily from request forms and through verbal inquiries made to the patient. Specifically, participants highlighted the relevance of these methods in gathering essential information such as the patient's age. This uniformity in responses suggests a commonality in the strategies employed by practitioners to populate DICOM metadata fields, emphasizing the significance of request forms and verbal interactions with patients as key information sources in this context.

5.2.3 Essential Metadata Fields for Examinations and Implications of Missing Metadata Data

The interview results explain the perceived cruciality of specific metadata fields for examinations, as well as the challenges faced when these fields are not provided. Participant 1 emphasized the significance of "Patient Name," "Patient ID," "Patient Age," "Gender," and "Clinical Details" as crucial metadata fields. In the absence of this information, obtaining details becomes very difficult, prompting the participant to send the patient back to the referring physician for completion. Similarly, Participant 2 identified "Clinical Details," "Examination," and "Patient ID" as critical fields, highlighting the practice of redirecting patients to the reception or referring physician when these essential details are missing.

5.2.4 Sequencing and Significance of DICOM Metadata Entry Across Examination Phases

The interview results provide insights into the sequence of entering DICOM metadata fields. Participant 1 outlined a systematic approach, noting that before the examination, essential information such as "Patient ID," "Patient Name," "Patient Age," "Gender," and "Examination" details are entered. During the examination, these fields may be further completed, and post-examination, the focus shifts to entering information related to labels. Similarly, Participant 2 described a pre-examination entry of "Patient ID," "Patient Name," "Patient Age," "Gender," and "Examination" details, with post-examination data entry concentrating on labels.

5.2.5 Determining Essential DICOM Metadata Fields for Examinations

The findings from the interviews show a shared viewpoint among participants concerning the importance of DICOM metadata fields completed post-examination. Both Participant 1 and Participant 2 expressed that fields completed post-examination are not considered less important. This consensus emphasizes the acknowledgment by practitioners that the timing of data entry does not diminish the importance of specific DICOM metadata fields, highlighting a recognition of the continued relevance and value of comprehensive data, regardless of when it is inputted in the workflow.

5.2.6 How to decide on the necessary DICOM metadata fields for a specific examination

The interview results indicate a consistent method employed by participants in determining the required DICOM metadata fields for a particular examination. Participant 1 and Participant 2 identified the request form as the primary source of information for establishing which DICOM metadata fields are necessary. This uniformity in responses underscores the instrumental role of request forms in guiding practitioners in the selection of essential metadata fields, highlighting a standardized and practical approach to data requirements in the context of medical examinations.

5.2.7 Instances of Omission of DICOM Metadata

The results show that in emergencies, neither participant fills in DICOM metadata fields before an examination. Instead, they both follow up after the examination to enter the details. This means that during urgent situations, providing immediate patient care is prioritized over entering metadata, but participants ensure to catch up on documenting the necessary information afterward.

5.2.8 Addressing Challenges in DICOM Metadata Entry

The interview results highlight varying experiences regarding challenges in entering DICOM metadata among participants. Participant 1 identified issues arising from unclear handwriting on request forms, necessitating additional questions to the patient for clarification, especially regarding clinical details. Furthermore, the participant mentioned verifying the patient's understanding of the examination instructions from the doctor before entering information. In contrast, Participant 2 reported no difficulties encountered during the data entry process. These findings underscore the significance of clear communication and patient cooperation in addressing challenges associated with DICOM metadata entry.

5.2.9 Policy Framework for Critical DICOM Tags

The interview results reveal insights into the existence and clarity of policies guiding the determination of important DICOM metadata fields. Participant 1 noted that clear policies are in place, particularly for cases involving age determination. Supervisors verbally communicated specific metadata fields—namely, patient name, age, and gender—that should be entered during examinations. Participant 2 similarly emphasized the clarity of policies in cases related to age determination.

5.3. Discussion of Medical Image Metadata Compliance Assessment

5.3.1 The University Teaching Hospitals

In the analysis of DICOM files from University Teaching Hospitals, the results revealed consistent completeness in essential patient identification fields across different years and modalities. However, notable gaps were identified in metadata fields such as Operators Name, Referring Physicians Name, and Institution, indicating potential areas for improvement in data entry practices. The decline in completeness for certain fields over the years, particularly in Operators Name and Referring Physicians Name, underscores the need for comprehensive data capture practices.

5.3.1 Levy Mwanawasa University Teaching Hospital

The results for Levy Mwanawasa UTH showed remarkable completeness across all fields in Digital Radiography - 2020, highlighting robust data capture practices. However, challenges were observed in Referring Physicians Name and Operators Name completeness in subsequent years. These findings suggest a need for further investigation into data entry processes and potential improvements.

5.4 Discussion of Results of Interviews with Radiographers

5.4.1 Additional information entered during an examination using modality machines

Interview results indicated commonalities and variations in the additional information entered during examinations, emphasizing the importance of "Examination" data. The differences in the sequence of entering "Patient Age" and "Patient Sex" suggest variability in data entry practices among participants.

5.4.2 Examining Information Sources

Consistency in the sourcing of information from request forms and verbal inquiries highlighted the significance of these methods in gathering essential information for DICOM metadata fields. This uniformity underscores the practical approach practitioners employ in populating DICOM metadata fields.

5.4.3 Essential Metadata Fields for Examinations and Implications of Missing Metadata Data

Practitioners emphasized the cruciality of specific metadata fields for examinations, such as "Patient Name," "Patient ID," and "Clinical Details." The challenges faced in the absence of this information underscore the importance of complete data for obtaining accurate details and patient care.

5.4.4 Sequencing and Significance of DICOM Metadata Entry Across Examination Phases

Insights into the sequence of entering DICOM metadata fields highlighted a systematic approach, with pre-examination and post-examination phases. This underscores the acknowledgment by practitioners that timing does not diminish the importance of specific DICOM metadata fields.

5.4.5 Determining Essential DICOM Metadata Fields for Examinations

The consensus among participants on the importance of post-examination completeness emphasizes the continued relevance and value of comprehensive data, regardless of when it is inputted in the workflow.

5.4.6 How to decide on the necessary DICOM metadata fields for a specific examination

The reliance on request forms as the primary source of information for determining necessary DICOM metadata fields indicates a standardized and practical approach to data requirements in medical examinations.

5.4.7 Instances of Omission of DICOM Metadata

During emergencies, participants prioritize immediate patient care over entering metadata, emphasizing the need for catch-up procedures afterward. This reflects a practical approach to balancing urgent patient needs with documentation requirements.

5.4.8 Addressing Challenges in DICOM Metadata Entry

Varying experiences in challenges highlight the significance of clear communication and patient cooperation in addressing issues associated with DICOM metadata entry. The need for clarity in handwriting on request forms and patient understanding of examination instructions is evident.

5.4.9 Policy Framework for Critical DICOM Tags

Insights into the existence and clarity of policies guiding important DICOM metadata fields reveal a structured approach, with verbal communication of specific metadata fields during examinations. This emphasizes the role of policies in guiding practitioners and ensuring standardization.

5.3 Chapter Summary

The examination of DICOM metadata practices in University Teaching Hospitals and Levy Mwanawasa University Teaching Hospital reveals both strengths and areas for improvement. Key findings include consistent completeness in essential fields, such as patient identification, while certain metadata fields exhibit lower completion rates, suggesting room for enhancement. Insights from interviews underscore the importance of clear communication, especially during emergencies. Additionally, the acknowledgment of policy frameworks highlights the need for standardization. Moving forward, addressing challenges and reinforcing policies will contribute to more robust DICOM metadata practices in medical imaging.

CHAPTER SIX

RECOMMENDATIONS AND CONCLUSION

6.1 Recommendations

Based on the results of the comprehensive analysis of medical image DICOM metadata from the University Teaching Hospitals (UTH) and Levy Mwanawasa University Teaching Hospitals, the following recommendations are proposed to improve the compliance of medical image metadata to the DICOM standard:

6.1.1 Enhance Data Entry Practices

Address the consistently low completeness percentages in critical fields such as Operators Name, Referring Physicians Name, and Institution across multiple years and imaging modalities. Implement training programs and standardized procedures to improve data entry practices for these fields [30].

6.1.2 Continuous Monitoring and Improvement

Continuous Monitoring and Improvement are important in ensuring metadata completeness and data entry practices. Establishing a systematic framework for ongoing evaluation allows for real-time assessment of DICOM fields, enabling the identification of patterns and areas for enhancement. Regular reviews, coupled with adaptive updates to training programs, ensure that personnel remain well-versed in the latest technologies and methodologies [30]. By fostering a culture of continuous improvement, healthcare institutions can proactively address challenges, optimize data entry processes, and consistently elevate the quality of metadata. This dynamic approach contributes to a more accurate and comprehensive representation of medical images, aligning with the evolving landscape of standards and technologies in the field.

6.1.3 Policy Refinement

Collaborate with relevant stakeholders to refine existing policies or establish new ones that explicitly outline the importance of each DICOM metadata field. Ensure these policies are communicated clearly and consistently to all practitioners.

6.1.4 Standardized Data Entry Procedures

Develop and enforce standardized procedures for entering DICOM metadata fields across all modalities and institutions. This includes ensuring consistent data entry practices for critical fields and addressing specific challenges identified during the assessment.

6.2. Conclusion

In conclusion, this research provides a systematic analysis of Medical Image Metadata, with a specific emphasis on compliance with the DICOM standard. It explores the organisation of diverse medical image data, considers ethical dimensions, and investigates metadata extraction challenges. The findings presented in this paper contribute valuable insights aimed at improving the overall adherence to metadata standards in medical imaging practices.

REFERENCES

[1] "Innovation Process in Medical Imaging," *Procedia - Social and Behavioral Sciences*, vol. 81, pp. 60–64, Jun. 2013.

[2] M. Aiello, G. Esposito, G. Pagliari, P. Borrelli, V. Brancato, and M. Salvatore, "How does DICOM support big data management? Investigating its use in medical imaging community," *Insights Imaging*, vol. 12, no. 1, p. 164, Nov. 2021.

[3] O. S. Pianykh, *Digital Imaging and Communications in Medicine (DICOM): A Practical Introduction and Survival Guide*. Springer Science & Business Media, 2009.

[4] "AAPM Reports - The Measurement, Reporting, and Management of Radiation Dose in CT." https://doi.org/10.37206/97 (accessed Dec. 04, 2023).

[5] "Website." http://dx.doi.org/10.1088/0031-9155/50/17/005

[6] "Science," *AAAS*. https://doi.org/10.1126%2Fscisignal.2001984 (accessed Dec. 04, 2023).

[7] K. Hellén-Halme, M. Nilsson, and A. Petersson, "Effect of monitors on approximal caries detection in digital radiographs—standard versus precalibrated DICOM part 14 displays: An in vitro study," *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.*, vol. 107, no. 5, pp. 716–720, May 2009.

[8] F. J. McEvoy and E. Svalastoga, "Security of Patient and Study Data Associated with DICOM Images when Transferred Using Compact Disc Media," *J. Digit. Imaging*, vol. 22, no. 1, pp. 65–70, Aug. 2007.

[9] R. Boellaard, "Standards for PET Image Acquisition and Quantitative Data Analysis," *J. Nucl. Med.*, vol. 50, no. Suppl 1, p. 11S–20S, May 2009.

[10] O. Bodenreider, "The Unified Medical Language System (UMLS): integrating biomedical terminology," *Nucleic Acids Res.*, vol. 32, no. suppl_1, pp. D267–D270, Jan. 2004.

[11] R. R. Beichel *et al.*, "Semiautomated segmentation of head and neck cancers in 18F-FDG PET scans: A just-enough-interaction approach," *Med. Phys.*, vol. 43, no. 6Part1, pp. 2948–2964, Jun. 2016.

[12] "Website." http://dx.doi.org/10.1007/s10278-010-9290-9

[13] "DICOM," *DICOM*. www.dicomstandard.org/ (accessed Dec. 04, 2023).

[14] M. Wada, Z. Ge, S. J. Gilmore, and V. J. Mar, "Use of artificial intelligence in skin cancer diagnosis and management," *Med. J. Aust.*, vol. 213, no. 6, pp. 256–259.e1, Sep. 2020.

[15] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection." https://doi.org/10.1109/ICCV.2017.324 (accessed Dec. 04, 2023).

[16] J. Xie, R. Girshick, and A. Farhadi, "Unsupervised Deep Embedding for Clustering Analysis," Nov. 19, 2015. Accessed: Dec. 04, 2023. [Online]. Available: http://arxiv.org/abs/1511.06335

[17] T. L. Morgan, "Quality Assurance for PET and PET/CT Systems," *Health Phys.*, vol. 103, no. 6, p. 810, Dec. 2012.

[18] S.-Y. Hu *et al.*, "Self-Supervised Pretraining with DICOM metadata in Ultrasound Imaging," in *Machine Learning for Healthcare Conference*, PMLR, Sep. 2020, pp. 732–749.

[19] A. Gosain and S. Sardana, "Handling class imbalance problem using oversampling techniques: A review." https://doi.org/10.1109/ICACCI.2017.8125820 (accessed Dec. 04, 2023).

[20] S. Venkataraman, "Digitization in Radiology: How Digital Tools are Converting Challenges into Opportunities," *CARRE4*, Dec. 13, 2020. https://medium.com/carre4/digitization-in-radiology-how-digital-tools-are-converting-challenges-into-opportunities-8b59ca2e248b (accessed Dec. 04, 2023).

[21] "Joblib: running Python functions as pipeline jobs — joblib 1.3.2 documentation." https://joblib.readthedocs.io/en/stable/ (accessed Dec. 04, 2023).

[22] "PySpark Overview — PySpark 3.5.0 documentation." https://spark.apache.org/docs/latest/api/python/ (accessed Dec. 04, 2023).

[23] M. Rocklin, "Dask: Parallel Computation with Blocked algorithms and Task Scheduling," in *Proceedings of the 14th Python in Science Conference*, 2015, pp. 126–132.

[24] M. M. Frost Jr, J. C. Honeyman, and E. V. Staab, "Image archival technologies," *Radiographics*, Mar. 1992, doi: 10.1148/radiographics.12.2.1561423.

[25] A. Bizzego *et al.*, "Evaluating reproducibility of AI algorithms in digital pathology with

DAPPER," *PLoS Comput. Biol.*, vol. 15, no. 3, p. e1006269, Mar. 2019.
[26] J. J. Alnasir, "Fifteen quick tips for success with HPC, i.e., responsibly BASHing that Linux cluster," *PLoS Comput. Biol.*, vol. 17, no. 8, p. e1009207, Aug. 2021.
[27] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, no. 6, pp. 305–311, Jun. 2020.
[28] S. T. Padmapriya and S. Parthasarathy, "Ethical Data Collection for Medical Image Analysis: a Structured Approach," *Asian Bioeth Rev*, pp. 1–14, Apr. 2023.
[29] "EscapeE." https://www.pclviewer.com/help/index.html?required_dicom_tags.htm (accessed Dec. 11, 2023).
[30] J. T. Norweck *et al.*, "ACR–AAPM–SIIM Technical Standard for Electronic Practice of Medical Imaging," *J. Digit. Imaging*, vol. 26, no. 1, pp. 38–52, Sep. 2012.

APPENDICES

Appendix A : Ethical Clearance Approval

UNIVERSITY OF ZAMBIA

BIOMEDICAL RESEARCH ETHICS COMMITTEE

Telephone: 260-1-256067
Telegrams: UNZA, LUSAKA
Telex: UNZALU ZA 44370
Fax: + 260-1-250753
Federal Assurance No. FWA00000338

Ridgeway Campus
P.O. Box 50110
Lusaka, Zambia
E-mail: unzarec@unza.zm
IRB00001131 of IORG0000774

11th May, 2023.

Ref. No. 2731-2022.

Dr. Ernest Obbie Zulu,
University Teaching Hospitals,
Adult Hospital Department of Radiology,
P/Bag RW 1X,
Ridgeway,
Lusaka.

Dear Dr. Zulu,

RE: APPLICATION FOR RENEWAL OF ETHICAL CLEARANCE APPROVAL: "ENTERPRISE MEDICAL IMAGING FOR STREAMLINED RADIOLOGICAL DIAGNOSIS IN ZAMBIAN PUBLIC HEALTH FACILITIES" (REF. NO. 2731-2022)

We acknowledge receipt of the request for renewal to the aforementioned study.

Renewal is hereby given for a period of one year from 5th May 2023 to 4th May 2024.

Yours sincerely,

Sody Mweetwa Munsaka, BSc., MSc., PhD
CHAIRPERSON
Tel: +26099925304
E-Mail: s.munsaka@unza.zm

Figure A.1: UNZABREC Ethical Clearance Form

NATIONAL HEALTH RESEARCH AUTHORITY

Paediatric Centre of Excellence, University Teaching Hospital, P.O. Box 30075, LUSAKA

Chalala Office Lot No. 18961/M, Off Kasama Road, P.O. Box 30075, LUSAKA

Tell: +260211 250309 | Email: znhrasec@nhra.org.zm | www.nhra.org.zm

Ref No: NHRA000024/10/05/2022 **Date: 10th May, 2022**

The Principal Investigator,
Ernest Obbie Zulu,
University of Zambia
Lusaka, Zambia.

Dear Ernest Obbie Zulu,

Re: Request for Authority to Conduct Research

The National Health Research Authority is in receipt of your request for authority to conduct research titled **"Enterprise Medical Imaging for Streamlined Radiological Diagnosis in Zambian Public Health Facilities."**

I wish to inform you that following submission of your request to the Authority, our review of the same and in view of the ethical clearance, this study has been **approved** on condition that:

1. The relevant Provincial and District Medical Officers where the study is being conducted are fully appraised;
2. Progress updates are provided to NHRA quarterly from the date of commencement of the study;
3. The final study report is cleared by the NHRA before any publication or dissemination within or outside the country;
4. After clearance for publication or dissemination by the NHRA, the final study report is shared with all relevant Provincial and District Directors of Health where the study was being conducted, University leadership, and all key respondents.

Yours sincerely,

Prof. Godfrey Biemba
Director/CEO
National Health Research Authority

Figure A.2: NRHA Ethical Clearance Approval

All Correspondence should be addressed to the Permanent Secretary
Telephone: +260 211 253040/5
Fax: +260 211 253344

In reply please quote:
MOH/
No.............................

REPUBLIC OF ZAMBIA

MINISTRY OF HEALTH

NDEKE HOUSE
P. O. BOX 30205
LUSAKA

16th May, 2022

Obbie Zulu
LUSAKA

RE: REQUEST FOR AUTHORITY TO CONDUCT RESEARCH

Reference is made to your letter dated 25th April, 2022 in which you requested the Ministry, for permission to conduct a research titled "***Enterprise Medical Imaging for Streamlined Radiological Diagnosis in Zambia Public Health Facilities***". I wish to inform you that my office has no objection to this request provided that;

1. The relevant Institution Director where the study is being conducted are fully appraised;
2. The final study report is cleared by NHRA before any publication or dissemination within or outside the country; and
3. After clearance for publication or dissemination by NHZRA, the final study report is shared with the Ministry.

Kindly ensure minimum interruption in health service delivery to the selected health you will undertake your research.

By copy of this letter, the Provincial, District Health Offices and facilities are advised to allow you undertake the above mentioned research and provide you with the relevant support.

Yours faithfully

Prof. Lackson Kasonka
Permanent Secretary- Technical Services
MINISTRY OF HEALTH

Figure A.3: Ministry Of Health Ethical Clearance Form

REPUBLIC OF ZAMBIA
MINISTRY OF HEALTH
University Teaching Hospitals -Adult

P/Bag Rw 1X
Fax: +260 211 250305 Lusaka - Zambia
e-mail: mduth@yahoo.com Tel: +260 211 253947 (Switch Board)
+260 211 251451

OFFICE OF THE SENIOR MEDICAL SUPERINTENDENT

Our Ref:

Your Ref:

5th September, 2022

Dr. Ernest Obbie Zulu
University of Zambia
Department of Library & Information Science
P O Box 50110
LUSAKA

Dear Dr. Zulu,

RE: REQUEST FOR AUTHORITY TO CONDUCT RESEARCH

The University Teaching Hospital – Adult is in receipt of your letter dated 5th September, 2022 in which you had requested to conduct a research titled ***"Enterprise Medical Imaging for Streamlined Radiological Diagnosis in Zambia Public Health Facilities" at the University Teaching Hospital."***

I wish to inform you that permission has been granted and you are advised to liaise with the Head of Department.

Yours faithfully,

Dr. Mwila Lupasha
Head Clinical Care
for/**Senior Medical Superintendent**
UNIVERSITY TEACHING HOSPITALS - ADULT

Figure A.4: UTH's Ethical Clearance Form

Radiology Department
University Teaching Hospitals
Private Bag RW 1X
Ridgeway
Lusaka, Zambia, 10101

Approved
5.09.2022

The Head Clinical Care (HCC)
University Teaching Hospitals
Private Bag RW 1X
Ridgeway
Lusaka, Zambia, 10101

September 2, 2022

REF: REQUEST FOR ACCESS TO PATIENTS' OLD RADIOLOGY REPORTS AND IMAGES FOR THE PURPOSE OF SYSTEMATICALLY ORGANIZING THEM TO FACILITATE THEIR EFFICIENT STORAGE AND RETRIEVAL AS PART OF A PILOT RESEARCH PROJECT — ENTERPRISE MEDICAL IMAGING FOR STREAMLINED RADIOLOGICAL DIAGNOSIS IN ZAMBIAN PUBLIC HEALTH FACILITIES.

Dear Sir,

I am a fourth year Specialty Training Programme (STP) student in Radiology (ID number STPRAD 19010107) at the UTHs Adult hospital. I am collaborating with Dr. Lighton Phiri from the University of Zambia in conducting the pilot project mentioned above under the UNZA DRGS. We hereby request for access to patients' old radiology reports and images stored in the library/department at this hospital. The UTHs have been deliberately selected for piloting of this project, which was approved by UNZABREC and NHRA (References: UNZA-2731-2022 and NHRA000024/10/05/2022).

The accessed data (images and reports) will be kept confidential before, during and after the digitisation process and the final package shall be handed over to the Head of Department, Radiology.

I look forward to hearing from you.

Ernest Obbie Zulu, MBChB
Mobile: +26 097 7 199434; Email: obbiernest@gmail.com

Figure A.5: Head of Clinical Care Clearance Form

Appendix B : Questionnaire

The University of Zambia

Department of Computer Science

Evaluating DICOM Compliance for Medical Images in Public Health Facilities in Zambia

Elijah Chileshe (Student ID: 2023027016)

PG-Dip. Computer Science

For more information or any queries, kindly get in touch on 0973297682

Dear Respondent,

My name is Elijah Chileshe, a student at the University of Zambia, conducting a study on ***"Evaluating DICOM Compliance for Medical Images in Public Health Facilities in Zambia."*** This study focuses on assessing the level of compliance with DICOM (Digital Imaging and Communications in Medicine) standards within the realm of medical imaging. The evaluation seeks to determine the degree of adherence to these standards, playing a pivotal role in enhancing interoperability and consistency in medical image-generated data.

The purpose of this interview is to gather essential information regarding the necessary DICOM metadata fields that Radiographers are required to complete when conducting a patient examination. We are focusing on analyzing DICOM tags, which serve as the information radiologists examine. Our interest lies in understanding the specifics associated with medical image metadata and, particularly, in understanding how you provide descriptive information for images in the examinations you perform.

Your cooperation will be greatly appreciated.

For more information or any queries, kindly get in touch with the following:

Principal Investigator
Elijah Chileshe
0979718410
elijah.chileshe@cs.unza.zm
Supervisors
Mr. Claytone Sikasote,
Computer Science Department,
University of Zambia.
claytone.sikasote@cs.unza.zm,
Dr. Lighton Phiri,
Department of Library and Information Science,
University of Zambia.
lighton.phiri@unza.zm

Part 1:

INTERVIEW GUIDE QUESTIONS
1. When carrying out an examination using the modality machines, in addition to entering the patient ID, patient name, and patient date of birth, what other specific information do you enter? 2. Where do you source the information for DICOM metadata fields? (e.g., forms) 3. Which metadata fields are crucial for an examination? a. If the crucial metadata fields are not provided, how difficult is it for you to obtain that information? 4. Can you describe the order of entering DICOM metadata fields? Are certain fields completed before, during, or after the examination? a. What fields are entered before, during and after the examination? 5. If some fields are filled in after the examination, are they considered to be less important? 6. How do you determine which DICOM metadata fields are required for a particular examination? 7. Are there instances where you choose not to specify certain DICOM metadata fields? If so, how are these decisions made? 8. Are there any challenges or difficulties you encounter when entering DICOM metadata, and how are these addressed within your workflow? 9. Is there a policy in place which specifies which tags are important?

Part 2:

Kindly check/tick the DICOM Metadata fields that you enter when carrying out an examination on a patient:

- ☐ (0010,0010) Patient's Name
- ☐ (0010,0020) Patient ID
- ☐ (0010,0040) Patient's Sex
- ☐ (0010,0030) Patient's date of birth
- ☐ (0010,0032) Patients birth time
- ☐ (0010,1010) Patient's Age
- ☐ (0020,0020) Patient orientation
- ☐ (0020,0010) Study ID
- ☐ (0008,0020) Study Date
- ☐ (0008,1070) Operators' Name
- ☐ (0008,0030) Study time
- ☐ (0008,0050) Accession Number
- ☐ (0008,0090) Referring Physician's Name
- ☐ (0020,0011) Series number

DICOM Metadata Elements

Limited DICOM attributes | All DICOM attributes

Search:

Patient

Patient's Name: **4932404**
Patient ID: **4932404**

Station

Manufacturer: **Canon Inc.**
Manufacturer's Model Name: **CXDI**
Station Name: **CXDI**

Study

Study Instance UID: **1.2.392.200046.100.2.1.176694098434843.140116124616**
Study Date: **16 Jan 2014**
Study Time: **12:46:16pm**
Study ID: **3974**

Series

Series Instance UID: **1.2.392.200046.100.2.1.176694098434843.140116124616.1**
Series Date: **16 Jan 2014**
Series Time: **12:46:17pm**
Series Number: **1**
Modality: **CR**
Series Description: **pa**
Body Part Examined: **CHEST**

DICOM Object

SOP Instance UID: **1.2.392.200046.100.2.1.176694098434843.140116124616.1.1.1**
SOP Class UID: **1.2.840.10008.5.1.4.1.1.1**
Transfer Syntax UID: **1.2.840.10008.1.2.1**
Instance Number: **1**
Laterality: **L**
Photometric Interpretation: **MONOCHROME1**
Samples per Pixel: **1**
Pixel Representation: **0**
Columns: **2688**
Rows: **2208**
Bits Allocated: **16**
Bits Stored: **12**

Image Acquisition

KVP: **120.0**