APPLICATION OF DATA DRIVEN DECISION MAKING IN CASE OF ZEWDITU MEMORIAL HOSPITAL ANTIRETROVIRAL THERAPY CLIENT MANAGEMENT

THESIS SUBMITTED TO ADDIS ABABA UNIVERSITY, COLLEGE OF BUSINESS & ECONOMICS, IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS IN BUSINESS ADMINISTRATION (MBA)

By: Mamo Edris

ID No.: GSD/3048/11

Advisor: Dr. Mesfin Fikre (PhD)

January, 2024

Addis Ababa, Ethiopia

Declaration

I, Mamo Edris, herby declare that this thesis entitled ***"Application of Data Driven Decision Making in case of Zewditu Memorial Hospital Antiretroviral Therapy Client Management"*** is my own original work except for quotations and citations which have been duly acknowledged, and that this document has not been submitted for a degree in any other universities.

Mamo Edris _____ _____

Signature Date

Statement of Certificate

This is to certify that this thesis entitled "***Application of Data Driven Decision Making in case of Zewditu Memorial Hospital Antiretroviral Therapy Client Management***" was undertaken by Mamo Edris for the partial fulfillment of Masters of Business Administration at Addis Ababa University, is his original work and not submitted earlier for any degree either at this University or any other University.

Mesfin Fikre (Ph.D)	_____	_____
Name of Advisor	Signature	Date

APPROVAL SHEET

This is to certify that this thesis entitled "***Application of Data Driven Decision Making in case of Zewditu Memorial Hospital Antiretroviral Therapy Client Management***" submitted to Addis Ababa University, College of Posta Graduate for the award of the Degree of Master of Business Administration is a research work carried out by Mamo Edris, under my guidance and supervision.

Therefore, I hereby declare that no part of this thesis has been submitted to any other university or institutions for the award of any degree.

Submitted by:

Mamo Edris	_____	_____
	Signature	Date

Approved by :

Mesfin Fikre (Ph.D	_____	_____
Name of Advisor	Signature	Date
Baymot Tadesse (Ph.D)	_____	_____
Name of External Examiner	Signature	Date
Dejene Tulu (Ph.D	_____	_____
Name of Internal Examiner	Signature	Date
_____	_____	_____
Name of Department Head	Signature	Date

DEDICATION

This thesis is dedicate to my wife Habiba , my daughters Amreen, Mina, Sebrin and Kausar and my son Mahammed.

ACKNOWLEDGEMENT

First and foremost, I would like to thank my Advisor Dr. Mefin Fikre for his guidance, provision of crucial feedbacks and continues encouragement throughout the research work.

Also I would like to thank my friend Asresu Tariku for his assistance on HIV and healthcare related matters as subject matter expert and continuous encouragement he provided me throughout the research work.

Last but not least, I would like to thank my wife Habiba , my daughters Amreen, Mina, Sebrin and Kausar and my son Mahammed for being supportive and understanding throughout the research work.

Contents

Acronyms

BI	Business Intelligence
AACAHB	Addis Ababa City Administration
AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral Therapy
HAART	Highly Active Antiretroviral Therapy
HIV	Human Immunodeficiency Virus
EMR	Electronic Medical Record
FHAPCO	Federal HIV/AIDS Prevention and Control Secretariat Office
MOH	Ministry of Health
WHO	World Health Organization
KDD	Knowledge Discovery in Database
UNAIDS	United Nations Programme on HIV/AIDS
PLWHA	People Living With HIV/AIDS

List of Tables

List of Figures

ABSTRACT

In today's dynamic business landscape, leveraging big data analytics is essential for gaining competitive advantages. However, despite the abundance of data in the healthcare sector, particularly in developing nations like Ethiopia, there remains a notable gap in comprehensive data analysis and decision-making. This study focuses on analyzing the Antiretroviral Therapy (ART) client data from Zewditu Memorial Hospital to uncover hidden insights spanning several years. By employing a quantitative and exploratory research framework, the study integrates Business Intelligence and Data Mining methodologies to investigate the impact of data-driven approaches on patient care quality, resource optimization, and operational efficiency within ART client management. Utilizing visualization techniques through Power BI and predictive modeling with electronic patient information, the study yields concrete results that unveil patterns and trends, offering invaluable insights for predicting treatment outcomes. These findings are anticipated to not only advance academic understanding but also provide practical guidance for care providers, emphasizing the crucial role of data-driven decision-making in optimizing ART client management processes.

Keywords: Business Intelligence, Data Mining, Antiretroviral therapy

CHAPTER ONE

Introduction

1.1 Background

1.1.1 Business Intelligence

Business Intelligence (BI) is technology-driven process of collecting, analyzing, and presenting data to support better business decision-making. It involves diverse tools, technologies, and practices that allow organizations to transform raw data into actionable insights. These insights can help businesses understand their operations, customer behavior, market trends, and various other factors crucial for their success(Turban et al., 2019).

BI starts with the collection of data from various sources, both internal and external. These sources might include sales data, customer data, financial records, market research, and more. This data is then integrated and stored in a centralized data warehouse, which ensures consistency and reliability. The data collection is followed by data cleansing and transformation; Raw data is often messy and inconsistent BI tools clean and transform data to remove errors, duplications, and inconsistencies, making it suitable for analysis. This ensures that decision-makers are working with accurate and reliable information (Davis & Johnson, 2019). BI tools enable the analysis of cleaned and transformed data using diverse techniques such as querying, data mining, and statistical analysis. The results are presented through interactive reports and dashboards that can provide insights into various aspects of the business, including sales trends, operational efficiency, and customer behavior (Smith & Jones, 2020).

BI can incorporate predictive analytics models that use historical data to make forecasts and predictions. For example, it can predict future sales, customer churn, or demand for products. Business administrators can use these insights to plan for the future. BI also helps in optimizing resource allocation. It can also provide insights into inventory management, helping to reduce carrying costs. BI tools can assist in monitoring the competitive landscape. By analyzing market trends, customer sentiment, and the performance of competitors, businesses can make strategic decisions to maintain or enhance their competitive position. Furthermore, BI supports long-term strategic planning by providing historical and current data to inform decision-makers (Johnson et al., 2021).

1.1.2 Business Intelligence in Healthcare

Advancements in technology are continuously reshaping various aspects of our daily lives, including the field of healthcare. While healthcare has been a longstanding industry with centuries of history, modern medicine has significantly extended average lifespans. The ongoing innovations in technology are poised to make the healthcare experience more user-friendly and beneficial (Rikke et al.,2017).

Healthcare organizations are adopting analytics and reporting tools to drive data-driven insights and improve patient care. This approach has been called "clinical business intelligence" or "healthcare BI." Clinical organizations have used BI to collect, analyze, and present data related to patient care, operations, and financial aspects. It enables healthcare providers and administrators to make data-driven decisions that improve patient outcomes, optimize resource allocation, and enhance the overall efficiency of healthcare organization (Luigi et al.,2023).

In recent years, the healthcare industry has seen a significant transformation in the way patient treatment planning is approached. BI tools and analytics have played a pivotal role in this evolution. Business intelligence empowers healthcare professionals to gather and analyze extensive volumes of patient data, encompassing symptoms, medical history, and diagnostic information. This facilitates a more comprehensive and data-driven approach to the planning of treatment strategies. The utilization of BI systems in healthcare has been shown to improve decision-making processes by providing healthcare providers with real-time insights, personalized treatment options, and the ability to identify patterns and trends in patient symptoms(Fatimetou et al., 2019).

Healthcare organizations use clinical BI in many aspects of their operations and patient care processes, including optimizing the coordination and scheduling of healthcare staff based on patient volumes. With the integration of BI tools, healthcare organizations can access and analyze vast amounts of patient data, such as admissions, discharges, and historical patient volumes. This data-driven approach allows healthcare administrators to forecast patient demand, allocate resources efficiently, and make informed decisions when scheduling staff. By leveraging BI, hospitals and healthcare facilities can ensure that staffing levels align with fluctuating patient volumes, improving patient care, reducing wait times, and increasing operational efficiency. This

approach is not only cost-effective but also enhances the overall patient experience and quality of care (Parisa et al. ,2020).

By harnessing patient data, medical histories, and real-time information, BI systems can generate timely alerts for healthcare providers, indicating the specific services and treatments that patients require. This approach minimizes the need for multiple visits and ensures that patients receive more comprehensive and coordinated care during a single appointment. It not only improves the overall patient experience but also reduces healthcare costs, as it minimizes redundant procedures and tests. With BI-driven provider alerts, healthcare professionals can deliver more efficient and personalized care, ultimately enhancing the quality of healthcare delivery and patient outcomes. Furthermore, through data analytics and predictive modeling, BI systems can sift through vast amounts of patient information, clinical data, and historical records to pinpoint individuals who are at higher risk or in need of ongoing medical attention (Hsiao et al., 2016).

The other way in which BI tools can benefit the healthcare is creating and monitoring scorecards to track readmission rates, helping healthcare organizations maintain the quality of care and adhere to regulatory standards. Business Intelligence technologies enable hospitals to compile, analyze, and visualize patient data to create scorecards. These scorecards include key performance indicators related to readmission rates, allowing healthcare administrators to identify trends and patterns in patient readmissions. By continuously monitoring these scorecards, healthcare providers can implement data-driven strategies to reduce readmission rates, improve patient outcomes, and manage resources more effectively, all of which are critical in today's value-based healthcare environment (Krishna et al., 2017).

BI also plays a pivotal role in resource allocation and cost management within the healthcare sector. By harnessing the power of data analytics, BI tools enable healthcare organizations to make more informed decisions regarding resource allocation, which is essential for optimizing cost-effectiveness. Business Intelligence applications provide healthcare managers with the ability to track and analyze resource utilization, including staff allocation, equipment usage, and medication inventories. These insights help in making real-time adjustments, reducing waste, and ensuring that resources are allocated efficiently, ultimately leading to cost savings and improved financial sustainability within healthcare institutions (Cresswell and Sheikh, 2013).

1.2 Statement of the Problem

The Human Immunodeficiency Virus (HIV), responsible for causing Acquired Immunodeficiency Syndrome (AIDS), represents a significant global public health challenge, particularly impacting individuals in low and middle-income countries. AIDS stands as the second leading cause of death among adolescents in Africa. Since the initial detection of the HIV epidemic in Ethiopia in 1984, AIDS has claimed the lives of millions, persisting as a prominent public health, social, and economic concern [MOH, 2007]. According to HIV Related Estimates and Projections, in Ethiopia, the total number of individuals living with HIV in 2022 amounted to 610,350, encompassing 36,812 children. In the same year, the total number of deaths attributed to HIV/AIDS alone reached 11,322. Focusing on the specific case of Addis Ababa, the number of people living with HIV in 2022 was reported as 112,185, including 3,206 children, with a total of 1,594 individuals succumbing to HIV/AIDS within the same year in Addis Ababa alone (EPHI, 2022).

HIV/AIDS has had profound economic consequences globally, exerting significant strain on healthcare systems, reducing productivity, and exacerbating poverty. The epidemic has led to increased healthcare expenditures, loss of a productive workforce, and heightened social and economic disparities. In Ethiopia, the economic impact of HIV/AIDS has been substantial, affecting sectors such as agriculture, education, and labor. The costs incurred for healthcare provision and the loss of human capital have implications for the nation's development(World Bank, 2021).

Antiretroviral therapy is treatment of people infected with HIV virus using anti-HIV drugs. Among many factors complicating the successful delivery of ART care services at the facility levels, none adherence to Antiretroviral therapy remains the major factor. Non-adherence to Antiretroviral therapy refers to a situation where patients who have been prescribed Antiretroviral medications do not consistently and correctly take their prescribed medications as instructed by their healthcare providers. Failure to adhere to the prescribed antiretroviral medication is recognized as a significant threat to both individual treatment outcomes and the potential spread of drug-resistant viruses within the community (Hecht FM et al, 1998).

Zewditu Memorial Hospital’s ART Unit established for ART service delivery in 2000, has evolved into Ethiopia's largest hospital providing HIV care and treatment. By July 2021, the hospital had

served 14,181 individuals living with HIV, including 295 children (Zewditu Memorial Hospital, 2018). The selection of this hospital for the study is based on its extensive experience in ART services over the years, substantial client volume, and the availability of electronic client records spanning several years.

In the context of Zewditu Memorial Hospital's management of ART clients, several challenges exist that hinder the hospital's ability to provide optimal patient care, allocate resources efficiently, and operate with high levels of operational efficiency. One of the central challenges is the issue of non-adherence to ART among patients. The ART care practice at the hospital is, unform counseling service is given for every patient both newly HIV positive as well as patients on follow up thus investing a lot of energy and resource where it may not be necessary while depriving some section of the population of the attention or focus they need to thrive in ART treatment.

To address this pressing issue, the study's general objective is to investigate and assess the impact of BI applications on the management of Zewditu Memorial Hospital's ART clients. More specifically, the study will focus on the utilization of data mining techniques as a BI tool to predict patient adherence to ART regimens. Currently, the hospital lacks a proactive and data-driven approach to identify patients at risk of non-adherence and intervene in a timely manner. This knowledge gap hinders the hospital's ability to allocate resources effectively and enhance overall operational efficiency in the management of ART clients.

While numerous studies have explored the impact of HIV in Ethiopia, notable gaps persist. These include a predominant focus on the clinical facets of ART client management, overlooking the broader socioeconomic ramifications of HIV/AIDS on healthcare systems and economies. Additionally, there is a lack of integration between data mining techniques and Power BI for predicting patient adherence to ART regimens, insufficient utilization of electronic health records for data-driven decision-making, and a reliance on primary data collected solely from individual patients in most research endeavors. Addressing these gaps could lead to a more comprehensive understanding of the multifaceted challenges posed by HIV/AIDS and enhance strategies for effective management and mitigation

1.2.1 Research Questions

This study is guided by the following research questions:

- How the client information currently collected, stored, and utilized within the ART client management system of Zewditu Memorial Hospital?
- What are the available Business Intelligence tools, and how do they align with the specific needs of healthcare management?
- How can BI tools contribute to the analysis of adherence patterns and patient outcomes in the context of ART client management?
- What are the anticipated benefits and challenges of implementing BI applications in the management of ART clients at Zewditu Memorial Hospital?

1.3 Objective

1.3.1 General objective

The general objective of this study is to investigate and assess the impact of Business Intelligence applications on the management of Zewditu Memorial Hospital's ART clients, with a focus on improving patient care, resource allocation, and operational efficiency.

1.3.2 Specific objectives

To address the general objective of this study, the specific objectives are:

1. To analyze the current state of ART client management at Zewditu Memorial Hospital
2. Explore and select BI tools that are well-suited for managing ART client data and related healthcare processes
3. Evaluate how BI applications can help in identifying treatment trends, adherence patterns, and patient outcomes for ART clients.
4. To analyze the long-term benefits and challenges of implementing BI in ART client management

1.4 Significance of the study

ART client management is crucial for the well-being of individuals living with HIV. The study aims to evaluate how BI applications can enhance patient care. Understanding how BI tools can contribute to better patient outcomes and adherence is significant for the patients themselves, as it can lead to improved health and quality of life.

Effective resource allocation is essential for healthcare facilities, especially in a resource-constrained setting. The study's focus on how BI can optimize ART client management, has implications for cost-effectiveness and sustainability. Zewditu Memorial Hospital and similar institutions can benefit from cost savings and better utilization of resources.

BI applications enable data-driven decision-making. This study can emphasize the importance of using data analytics to make informed decisions in healthcare management. The findings can serve as a model for other healthcare facilities looking to leverage data for better decision support.

Improved ART client management not only benefits individual patients but also has broader public health implications. Effective HIV treatment and care can contribute to reducing the spread of the virus. Therefore, the study can have a positive impact on public health in the community served by Zewditu Memorial Hospital.

The study can contribute to the body of knowledge on the application of BI in healthcare management, particularly in the context of ART. This can be of interest to researchers and academics in the fields of healthcare management, information technology, and public health.

The research may provide insights that can inform healthcare policies and practices in Ethiopia and potentially in other countries with similar healthcare challenges. The study's findings can influence decision-making at the policy level.

In summary, this research is significant because it addresses critical issues in healthcare management, with a specific focus on ART client care at Zewditu Memorial Hospital. The findings can lead to improved patient outcomes, resource optimization, and operational efficiency, benefiting both the hospital and the broader community.

1.5 Scope and Limitation

1.5.1 Scope

The scope of this study is limited to Zewditu Memorial Hospital, chosen due to its status as the largest ART client provider in Addis Ababa. This selection was made based on constraints related to time and budget. However, it is acknowledged that a broader investigation involving multiple sites across the city or country could yield more comprehensive insights. Given adequate resources and time, expanding the study to include additional sites has the potential to enrich the findings and provide a more holistic understanding of ART client management practices

1.5.2 Limitation

The data extracted from Zewditu Memorial Hospital's Enhanced EMR-ART application has a serious completeness issue. Due to this the researcher is forced to reduce the overall records and variables significantly by excluding variables with significant NULL values especially variables with categorical values that cannot be refilled for missing values using statical software such as SPSS. This affected the overall performance of the model because some important variables that are believed to have a significate impact on ART adherence prediction were excluded due to missing value issues.

CHAPTER TWO

Literature Review

2.1 Theoretical Literature Review

2.1.1 Healthcare management

The landscape of healthcare delivery and patient situations undergo continual transformations, necessitating managers to continuously acquire new skills and competencies. Management in healthcare encompasses various abilities, including motivating staff, effective communication and negotiation with stakeholders, and adopting specific attitudes and behaviors that enhance staff discipline and performance. Additionally, managers must grasp the fundamental technical aspects of the services provided. Although training courses are effective, they may not always suffice to impart all the required skills for these competencies (WHO, 2017).

Effectiveness, efficiency, and equity are pivotal concepts in healthcare management. Effectiveness in the health sector is attained when managers appropriately select goals and ensure that their healthcare teams can accomplish them. Efficiency, on the other hand, gauges how well resources are utilized to achieve those goals. A healthcare system is deemed efficient when funds and materials are utilized judiciously with minimal wastage. Efficiency involves the correct execution of tasks, prudent resource utilization, and minimizing waste. Health inequalities stem from the inequitable distribution of resources and may be linked to factors such as low-income levels, housing, education, gender, geographical accessibility, and occasionally ethnicity (WHO, 2000).

Healthcare management is a multifaceted field that encompasses a wide range of activities, from coordinating patient care to overseeing healthcare facilities and optimizing resources. Healthcare management encompasses the strategic processes of planning, organizing, directing, and controlling resources to achieve organizational goals with efficiency and effectiveness. It plays a vital role in improving the quality of care, patient safety, and overall healthcare system performance (Buchbinder and Shanks, 2016).

Quality improvement is a fundamental aspect of healthcare management. Healthcare leaders are continuously seeking ways to enhance the quality of care. Quality improvement efforts focus on systematically assessing, monitoring, and improving processes to deliver better patient outcomes

and experiences. Initiatives such as the Plan-Do-Study-Act (PDSA) cycle and Six Sigma methodologies are employed to drive quality improvements in healthcare organizations (Boaden et al. , 2008).

Healthcare management faces numerous challenges, including resource constraints, rising costs, and the need to balance the triple aim of healthcare: improving patient experience, population health, and reducing costs. The challenges of managing healthcare organizations in the face of changing healthcare delivery models, emphasizing the need for innovative management strategies to navigate these complexities (Rundall et al. ,2013).

Leadership is a critical component of healthcare management, influencing the culture and success of healthcare organizations. The role of leadership in healthcare, stating that effective leaders create an environment where staff members are engaged, motivated, and aligned with organizational goals. Leadership theories, such as transformational leadership, have been applied to inspire positive change and innovation in healthcare management (Ginter, Duncan, and Swayne, 2018).

The future of healthcare management is shaped by ongoing advancements in technology, shifting demographics, and evolving patient expectations. Healthcare managers are increasingly embracing data analytics, telemedicine, and patient-centered care models to adapt to the changing landscape. The integration of technology and data-driven decision-making is poised to further enhance healthcare management in the years ahead (Wise, J., 2017).

2.1.2 Healthcare management in Ethiopia

Ethiopia has made significant progress in expanding access to ART for HIV patients. The government, with support from international organizations, has implemented programs to provide free ART to those in need. Tadele et al. (2017) conducted a study on the effectiveness of ART in Ethiopia and found that while progress has been made, challenges in patient adherence and retention in care persist. This underscores the importance of robust management systems and adherence support for ART clients.

The Ethiopian government has launched various initiatives to improve healthcare quality and management. The Health Sector Transformation Plan (HSTP) is a notable example, aiming to enhance the quality of healthcare services across the country. Teklegiorgis et al. (2020) assessed

the impact of the HSTP and found that it has led to improvements in healthcare infrastructure, staffing, and the overall quality of healthcare services. These improvements are critical for the effective management of healthcare, including ART client management.

While Ethiopia has made significant strides in expanding access to ART, challenges related to patient adherence and retention in care persist. Adane et al. (2020) conducted a study focusing on ART adherence in Ethiopia and identified factors such as medication side effects, stigma, and long travel distances to healthcare facilities as barriers to consistent adherence. Effective healthcare management strategies must address these challenges to ensure that ART clients receive continuous and quality care.

Building a skilled healthcare workforce is crucial for effective healthcare management in Ethiopia. The country has been working to improve healthcare education and training programs. A study by Kinfu et al. (2016) emphasizes the importance of building the capacity of healthcare workers, particularly in rural areas. Strengthening healthcare management and workforce capacity is essential to effectively manage the increasing number of ART clients.

Ethiopia has been exploring the use of technology, including telemedicine, to improve healthcare access and delivery. A study by Gebru et al. (2021) investigated the adoption and acceptance of telemedicine among healthcare professionals in Ethiopia. The findings suggest that telemedicine can be a valuable tool for managing ART clients in remote or underserved areas, where access to healthcare facilities may be limited.

2.1.3 Business Intelligence

Wieder et al. (2012) highlighted Business Intelligence as a prominent topic in the information technology industry. In today's banking sector, there is an increased acquisition of data with greater granularity and volume. Consequently, the primary challenge in information management has shifted from collecting sufficient data to the efficient analysis and utilization of this data (Chaudhuri et al., 2011). Enterprise data is often dispersed across various operating systems, and external information structures may lack compatibility and a clear definition of how they can be utilized, leading to a fragmented understanding of the subject matter (Hovi et al., 2009). Business Intelligence systems play a crucial role in integrating, analyzing, and accessing this dispersed

information throughout the organization, supporting both operational and strategic decision-making processes (Hovi et al., 2009).

Gilad and Gilad (1988) asserted that the organizational function of Business Intelligence (BI) is responsible for facilitating the development of BI. They emphasized that although BI can be viewed as an informal activity carried out by all employees who use information to fulfill their individual needs, an organizational role should be integrated and treated as a structured, organized operation closely tied to strategic management. Consequently, maintaining a consistent BI approach helps minimize data loss, reduce duplicates, and, most importantly, integrate data in various formats into a cohesive framework for strategic planning from diverse sources (Choo, 2002.

Ramakrishnan et al. (2012) elucidate the three overarching purposes for the implementation of Business Intelligence (BI). The first objective is for organizations to gain insight, as competitive pressures in the market increase complexity. The authors contend that BI systems are becoming essential for companies to navigate the progressively unpredictable business landscape. They introduce second and third BI values linked to the coherence of information within an entity. The authors posit that BI provides a singular version of truth for an organization and can also facilitate organizational transformation. Given that organizational data is in constant flux, especially with the growth of companies, having a singular version of truth enhances communication among employees when they all have access to the same information. The clear business representation of numbers, calculations, and expressions also improves data quality and saves time for more effective analysis (Ramakrishnan et al., 2012).

Business Intelligence applications have emerged as a transformative force in healthcare management, facilitating data-driven decision-making and improving patient care outcomes. BI tools in healthcare have evolved to efficiently process vast amounts of data from electronic health records, billing systems, and other sources. These tools provide healthcare administrators with the capability to generate comprehensive reports and dashboards, aiding in resource allocation, patient management, and financial analysis. The integration of BI systems enables healthcare professionals to make informed decisions based on real-time data, enhancing the quality of care and operational efficiency(Wager et al., 2017).

The application of Business Intelligence (BI) in healthcare management, particularly in the context of Antiretroviral Therapy (ART) client management, has demonstrated significant potential in enhancing patient care, resource allocation, and operational efficiency. Studies such as Wang et al. (2019) have shown that BI tools and practices, when integrated into healthcare systems, enable healthcare providers to monitor patient progress, adherence, and clinical outcomes more effectively. This not only leads to improved patient care but also empowers healthcare organizations to allocate resources more efficiently. By extracting actionable insights from large volumes of healthcare data, BI supports data-driven decision-making, contributing to better outcomes for ART clients and the overall effectiveness of healthcare delivery.

2.1.4 Data driven decision making

The contemporary healthcare sector produces extensive and intricate data encompassing patient information, hospital resources, disease diagnoses, electronic patient records, medical devices, and more. This substantial volume of data serves as a pivotal resource requiring processing and analysis to extract knowledge. This knowledge can then be utilized for informed decision-making, ultimately enhancing the efficiency and cost-effectiveness of healthcare services (Simsek et al., 2023).

During the provision of ART services to clients, a substantial amount of data is accumulated. Over time, this data holds valuable insights into the factors influencing the consistent utilization of ART care. However, as the data volume increases, it becomes laden with implicit knowledge that eludes simple statistical analysis methods. Additionally, the larger the database, the more challenging it becomes to leverage the data for intelligent decision support systems. Consequently, conventional query and report-based systems, along with statistical analysis techniques, prove inadequate for this purpose. The most suitable approach to uncover and manage the implicitly embedded knowledge within the collected data is through data mining (AbdulKareem, 2000).

Concerns regarding data overload and the absence of suitable analysis tools for generating decision support information in the healthcare system have been acknowledged in various literature. For instance, Kaur et al. (2006) articulated this challenge as, "The healthcare environment is generally perceived as being 'information rich' yet 'knowledge poor." They further elucidated, stating, "There is a wealth of data available within the healthcare systems. However, there is a lack of effective

analysis tools to discover hidden relationships and trends in data. Valuable knowledge can be discovered from the application of data mining techniques in the healthcare system."

Data mining amalgamates techniques from database research, artificial intelligence, and statistics, finding applications across various domains. The concept of data mining is gaining popularity as an information management tool, uncovering implicit knowledge within data collected by various application areas over time. Data mining methodologies can be employed in clinical databases to provide decision support. Clinical healthcare data is extensive, encompassing patient-centric, resource management, and transformed data. Healthcare organizations need the capability to analyze this vast amount of data. By computerizing and storing treatment records for millions of patients, data mining techniques can play a crucial role in addressing essential and critical questions related to healthcare (Fayyad et al., 1996).

In contemporary times, the term "big data" has gained widespread recognition and is commonly used to depict the substantial expansion, accessibility, and utilization of data or information. Laney's (2001) conceptualization of big data revolves around three primary characteristics known as the three V's: Volume, Variety, and Velocity. Big data is distinguished by the presence of large quantities of data sourced from diverse outlets, coupled with a rapid generation speed. This form of extensive data is seamlessly integrated into our daily routines, as individuals routinely interact with it indirectly through various activities, such as online searches or reading articles (EY, 2014).

When organizations actively incorporate and utilize data in their decision-making processes, they transition into becoming data-driven entities (Anderson, 2015). Moreover, in a fundamental sense, big data can be categorized into two main types: structured data and unstructured data (Turban, 2015; EY.com, 2014; Hand, 2007; Stone, 2014). Structured data refers to information that has already been organized, follows a predictable format, and is defined and governed by a set of rules. Conversely, unstructured data lacks predefined rules and regulations, resulting in an absence of a predictable format.

The capability to derive meaningful insights from unstructured data using analytics empowers organizations to integrate valuable knowledge about their business status into their daily operations, processes, and decision-making. This implies that every organization's data users must collect data, analyze it, transform it into insights to gain knowledge, and subsequently ensure that

the outcomes inform actions (Anderson, 2015). If executed effectively with high quality, organizations can obtain crucial competitive advantages by being well-informed about their business and leveraging this information to enhance their performance (Brynjolfsson, 2011; Marr, 2015).

The process of data-driven decision-making involves deciding on a course of action based on big data (Williams & Brunner, 2011). With the increasing accessibility and intelligence of data systems and technologies, utilizing big data to inform decision-making has become more feasible in various industries, including the telecom industry and other services. It goes beyond merely dealing with numbers or extracted data; it involves transforming data into actionable and usable knowledge.

2.1.5 Business Intelligence Tools

In the era of overwhelming data, managing and extracting meaningful insights has become a formidable challenge. The sheer volume makes it impractical for an individual to meticulously sift through data line by line to discern distinct patterns and observations. This is where data visualization becomes essential – providing a visual summary that significantly simplifies the identification of patterns and trends compared to scrolling through numerous rows on a spreadsheet. Aligned with the inherent workings of the human brain, visualized information is more intuitive and accessible. Given that the ultimate goal of data analysis is to gain insights, the value of data is greatly enhanced through visualization. Even if a data analyst can derive insights without visualization, effectively communicating the meaning becomes more challenging without the aid of visual representation (Akhtar et al., 2020).

Data visualization involves representing quantitative information through graphical means, converting both extensive and compact datasets into visual formats that are more comprehensible and digestible for the human brain. While familiar charts and graphs are commonplace in our daily lives, they play a particularly crucial role in business intelligence by aiding users in making informed decisions based on data. The process of data visualization serves to convert raw data into meaningful information, encompassing three primary subfields; Scientific Visualization, Information Visualization and Visual Analytics (Akhtar et al., 2020).

Data mining is a powerful Business Intelligence tool that enables organizations to extract valuable patterns, insights, and knowledge from large datasets. It involves the use of various algorithms and statistical techniques to uncover hidden relationships, trends, and anomalies in data, which can inform strategic decision-making. Data mining helps businesses gain a competitive edge by identifying customer preferences, market trends, and operational efficiencies (Han et al., 2011).

Data mining, also referred to as knowledge discovery in databases, is defined as the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data (Fayyad et al, 1996). It has garnered significant attention in the information industry and society as a whole in recent years, mainly due to the widespread availability of vast amounts of data and the pressing need to transform such data into valuable information and knowledge. The information and knowledge derived from data mining can find applications in diverse fields, including market analysis, fraud detection, customer retention, production control, scientific exploration, and healthcare (Han and Kamber, 2016).

As stated by the authors in (Hian and Tan, 2005), there is extensive potential for the application of data mining in the healthcare sector. Broadly, these applications can be categorized into the assessment of treatment efficacy, healthcare management, customer relationship management, and the identification of fraud and abuse.

As outlined by Milley A. (Milley, 2000), the utilization of data mining applications can be designed to assess the efficiency of medical treatments. Through the comparison of causes, symptoms, and treatment courses, data mining can provide an analysis of the effectiveness of different courses of action.

To support healthcare management, the development of data mining applications can enhance the identification and monitoring of chronic disease states and high-risk patients. It enables the design of suitable interventions and helps in minimizing hospital admissions and claims, as stated by (Biafore, 1999).

The increasing popularity of data mining in healthcare can be attributed to the generation of large, heterogeneous, and complex data in healthcare transactions (Ciosa and Mooree , 2002). This data is too intricate and voluminous for analysis through traditional statistical methods typically used for relatively small amounts of primary data collected with specific objectives (hypothesis testing and probabilistic inference) (Hand DJ, 1998). However, data mining, in general, has the capacity

to analyze and enhance decision-making by uncovering patterns and trends in extensive and complex secondary data (Ciosa and Mooree, 2002). The insights derived from data mining can impact cost, revenue, and operational efficiency while upholding a high level of care (Silver M et al, 2001). Consequently, data can serve as a valuable asset for healthcare organizations, but the initial transformation into useful information and knowledge is crucial (Han and Kamber , 2016).

2.2 Empirical Literature Review

A study by Chen et al. (2018) emphasizes the role of BI in healthcare resource allocation and operational efficiency. The authors found that BI applications are instrumental in optimizing staff scheduling, streamlining inventory management, and improving the overall functioning of healthcare organizations. The study highlights the cost-saving potential of BI in healthcare, which can be redirected toward patient care and other critical services. Additionally, BI helps in tracking performance metrics and identifying areas for improvement, contributing to the continuous enhancement of healthcare management.

Moreover, research by Kim et al. (2018) underscores the advantages of BI in patient care optimization. The study demonstrates that BI tools play a pivotal role in the analysis of patient data, enabling healthcare providers to identify treatment trends and adherence patterns. This allows for more personalized care, which can lead to better patient outcomes. The ability to monitor patient progress and promptly identify issues contributes to enhanced clinical decision-making, resulting in a significant improvement in patient care quality within healthcare management

Cao et al. (2015) investigated the impact of business analytics on decision-making effectiveness using a path model analysis. The study, based on survey responses from 740 employees in UK enterprises, revealed a positive influence of business analytics on decision-making effectiveness. The researchers found that business analytics, mediated by a data-driven environment, enhances information processing capabilities, ultimately contributing to improved decision-making within organizations.

Bayrak (2015) conducted a study reviewing the status of business analytics as either a business enabler or a transient trend. The research drew upon publicly accessible data from reputable databases such as sciencedirect.com, wiley.com, springer.com, emeraldinsight.com, tandfonline.com, informs.org, and dl.acm.org. The findings underscored a notable trend where an

increasing number of companies are turning to business analytics. These organizations leverage analytics to strategically plan and optimize their operations, project future business outcomes, enhance overall efficiency, make informed decisions, introduce innovative products and services, and seize emerging market opportunities.

Shanks and Sharma (2011) conducted a study on deriving value from business analytics systems, specifically exploring the influence of strategy. Employing an explanatory case study research approach, the researchers contended that dynamic business analytics capabilities serve as a critical tool for comprehending the mechanisms through which business analytics technology translates into value-driven initiatives and, consequently, enhances overall firm performance. The study further asserted that organizational strategy plays a pivotal role in shaping both the quality of business analytics technology and the organizational structure. Additionally, the researchers highlighted the impact of organizational strategy on fostering value-centric actions within the business analytics framework.

A research conducted in Botswana by Kip, Ehlers, and van der Wal (Lankowski et al., 2014) discovered that socioeconomic factors at the patient level, including unemployment, illiteracy, and a low level of education, can result in a limited understanding of medication effectiveness and contribute to decreased treatment adherence. The likelihood of non-adherence was notably elevated when patients faced challenges in reading and comprehending basic written medical instructions (Martin, et al., 2005)).

Data mining investigations in the context of administering antiretroviral treatment are intended to inform the treatment strategies for ART clients. A research study conducted in South Africa demonstrated the potential to predict changes in CD4 count using support vector machines (Singh Y, Mars M, 2010). The study achieved the highest accuracy of 83% by incorporating factors such as genome, current CD4 count, and the number of weeks from the baseline CD4 count. The ability to forecast CD4 cell count provides valuable assistance to clinicians in treatment management, resource allocation, and determining the most suitable therapy for individual patients.

Another investigation delves into the utilization of support vector machines (SVM) to forecast drug resistance in HIV strains derived from patients, relying on the genetic sequences of the viral DNA sections responsible for encoding the crucial enzymes, Reverse Transcriptase or Protease,

vital for HIV replication. The SVM model demonstrated an accuracy performance ranging between 94.13% and 96.33% (Seara T, 2009).

2.3 Conceptual Framework

Cross- Industry Standard Process for Data Mining (CRISP-DM) is adopted in this research as conceptual framework for Data Mining modeling as well as BI data preparation. CRISP-DM, is a popular framework that outlines the key steps and stages involved in the process of data mining and analytics (Wirth and Hipp, 2000). It provides a structured and systematic approach to guide organizations and data scientists through the complexities of data-driven projects. CRISP-DM is applicable across various industries and is particularly well-suited for developing predictive models and gaining insights from data (Ananda et al., 2022).

CRISP-DM outlines a structured approach for data mining projects, consisting of six phases. The process begins with Business Understanding, where project objectives and requirements are established from a business perspective. This is followed by Data Understanding, where data sources are identified, collected, and explored to gain insights into their quality and structure. Next is Data Preparation, involving activities such as cleaning, transforming, and selecting data to make it suitable for analysis. The fourth phase is Modeling, where various data mining techniques are applied to build and evaluate models that address the business objectives. Once satisfactory models are developed, they are deployed in the fifth phase, Evaluation, where their performance is assessed against the project's goals. Finally, the process concludes with Deployment, where the successful models are integrated into the business process, and plans are made for monitoring and maintenance(Chapman et al., 2000)..

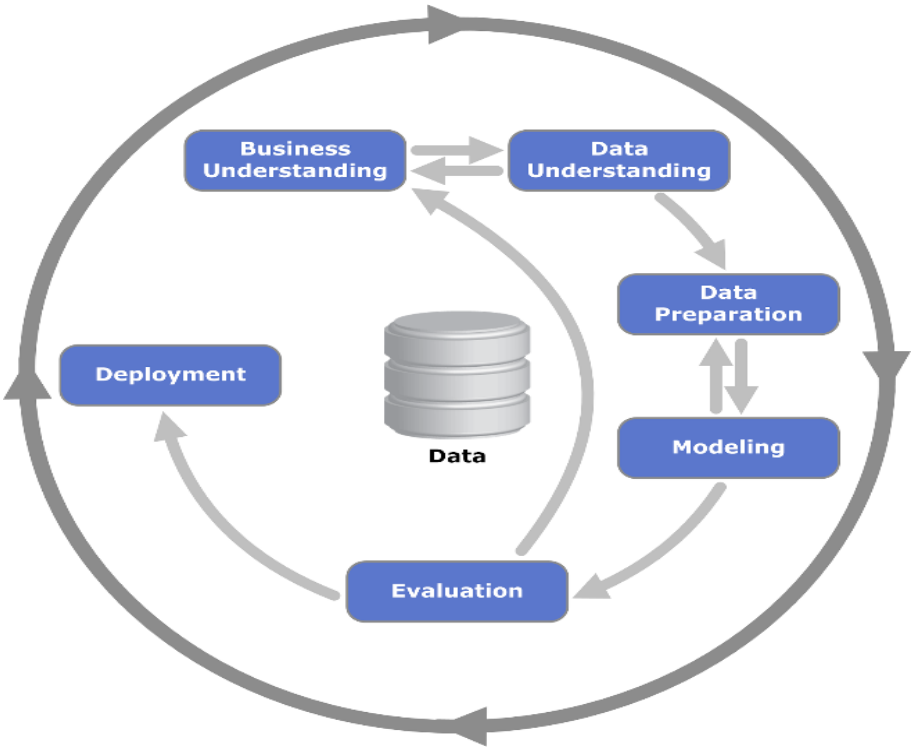

Figure 1: CRISP-DM Diagram

Business Understanding

As part of Business Understanding the researcher has made effort to understand how the hospital manages the ART patients and relevant resources. Also the entire process of the hospital's ART Clinic services explored patient intake, follow up, counseling, lost to follow up tracing and others.

ART is lifelong commitment, accompanied by substantial costs. It is an integral component of comprehensive care, encompassing services such as Voluntary Counseling and Testing (VCT), management of Sexually Transmitted Diseases (STDs), Tuberculosis (TB), Opportunistic Infections (OI), and the prevention of mother-to-child transmission (PMTCT), including the treatment of pregnant women in ART clinics within hospitals and health centers. The successful implementation of ART involves suppressing HIV viral replication, thereby decelerating disease progression, enhancing immunity, and extending overall life expectancy. However, ensuring adherence to ART poses a significant challenge in the clinical management of HIV. The primary interventions both before and after initiating therapy, along with fundamental services and procedures adhered to in clinics to guarantee client adherence, are outlined in accordance with WHO guidelines(WHO guideline).

As per the WHO guidelines, the Key Interventions to Enhance Adherence of Patients before Starting ART Include the Following (WHO, 2000):

- Assessing patient readiness for drug treatment.
- Identifying the types of support that will optimize the patient's adherence to therapy.
- Addressing the patient's individual learning needs, including the importance of ART adherence and the consequences of non-adherence.

There are also Interventions after Starting ART to Enhance Clients' Adherence. These include (WHO, 2000):

- Ongoing education about the importance of adherence and consequences of non-adherence.
- Social support to improve adherence.
- Use of reminders to take medications, such as medication charts
- Improved communication with healthcare workers.
- Continuous reinforcement of adherence.

Due to ART provision complexity, there is a defined series of procedures to be followed while providing the service which is part of the World Health Organization (WHO) guideline. Accordingly, doing first and second visit of the HIV patient the Medical Officer/ART Physician conducts a series of medical tests and gather relevant information from the patient and inform the patient of his/her status test results and determine if the patient is ready to start ART which is also based on the criterial stipulated on the WHO guideline.

After the ART service provider determined probable eligibility for ART and discussed treatment initiation, the ART Client is referred to the adherence counselor for counseling sessions. HIV patient counseling sessions adhere to the guidelines outlined by the WHO. These sessions play a crucial role in supporting individuals throughout their ART journey. The WHO ART guidelines emphasize a patient-centered approach, focusing on holistic care that goes beyond medical intervention. Counseling sessions cover various aspects, including treatment initiation, medication adherence, potential side effects, and the importance of consistent follow-up. Patients receive information tailored to their needs, ensuring comprehension and active participation in decision-making regarding their health. Furthermore, these sessions address psychosocial aspects, fostering emotional well-being, and providing a supportive space for patients to express concerns or seek

clarification. By aligning with WHO guidelines, HIV patient counseling sessions aim to enhance treatment adherence, empower individuals to actively engage in their healthcare, and contribute to overall improved health outcomes.

The ART Client meets with the ART adherence counselor after the client's medical visit every two weeks for first two months and once a month then after. After a duration of at least three months for adults and six months for children on the therapy, the patient will attend the monthly visit with the Triage Nurse. If the patient is considered "stable" on ART, indicating the absence of medication side effects or effective management of any existing side effects, and the patient demonstrates strong adherence by not missing a single ARV dose since the previous visit, the Triage Nurse will conduct a brief assessment of the patient's status and issue a prescription for the next month's supply of Antiretrovirals. The details of the patient's visit are recorded in the Follow-Up form.

With the standard ART clinic services and procedures outlined above, Zewditu Memorial Hospital ART clinic caters to all clients, regardless of their anticipated adherence category (Adherent or Non-Adherent). However, this approach results in unnecessary efforts being extended to clients who face no adherence issues after initiating ART, while those requiring additional support receive the same standard procedures. Consequently, reallocating this surplus effort from clients with no adherence problems to those in need of additional assistance contributes to an overall improvement in clients' adherence. Despite providing pre-ART counseling, the clinic follows up with and counsels clients at each drug picking appointment to ensure strict adherence. However, the clinic cannot tailor these efforts based on future adherence categories since there is no way to predict a client's category accurately without resorting to expert judgment. Hence, classifying clients into adherence categories based on pre-ART information before initiating therapy is crucial. This classification enables the application of category-specific efforts, thereby increasing adherence. Therefore, one of the research objectives is to construct a classification model using ART dataset, allowing the prediction of therapy adherence categories for new clients based on pre-ART attributes. This information aids professionals in tailoring specific counseling and follow-up strategies for each category, ultimately enhancing clients' adherence.

CHAPTER THREE

Methodology

3.1 Introduction

The preceding chapter has set the stage by providing the context for the literature review. This chapter shifts the focus towards elucidating the research methodology employed to achieve the desired objectives. The structure of this chapter is delineated as follows: research approach, research design, identification of data sources, data collection methods, and elucidation of the data analysis methods. The intention is to furnish both the researcher and readers with a comprehensive roadmap detailing how the research was conducted.

3.2 Research Approach

The research study aims to investigate and evaluate the utilization of data-driven decision-making processes in the management of ART clients at Zewditu Memorial Hospital. The study recognizes the pivotal role of data in healthcare decision-making and seeks to extract valuable insights from the hospital's HIV Patient database. By employing a quantitative research approach, the study focuses on extracting and analyzing pertinent data from the hospital's HIV Patient database, including information related to patient demographics, treatment adherence, health outcomes, and other relevant variables. The study aims to investigate how strategies guided by data-driven approaches can improve the quality of patient care, optimize the allocation of resources, and enhance overall operational efficiency within the realm of ART client management.

3.3 Research Design

Following exploratory research design, the research design aims to investigate the application of data driven decision making in the management of Antiretroviral Therapy clients at Zewditu Memorial Hospital. The approach integrates both Business Intelligence (BI) and Data Mining methodologies, focusing on the utilization of electronic HIV patient data even though the ART service provision practice is explored extensively by talking to subject matter experts and going through ART service provision guideline by WHO to have understanding of the current practice and identify the gap to explored through this research. The aim is to comprehensively explore how data-driven strategies can enhance decision-making processes in the healthcare setting, particularly in the context of ART client management. The BI component focuses on visualization while the

Data Mining component introduces predication of patient treatment outcome by utilizing the existing electronic patient information.

3.4 Data Source

The study encompasses the entire adult population of ART clients recorded at Zewditu Memorial Hospital, spanning from November 24th, 2000, to January 30th, 2022. The specific study population include all adult ART clients who have initiated ART services at Zewditu Memorial Hospital within the stipulated timeframe. However, certain records from this specified population excluded such as ever started adult ART clients whose information is incomplete and whose baseline information & class attribute value not known.

3.5 Study Variables

Dependent variable/ Class Attribute: Adherence to ART

Independent Variables:

Socio-demographic characteristics: - Sex, Age, Marital Status, Level of Education, Religion, ART initiation Site, and Region.

Baseline characteristics: - WHO clinical staging, functional status, Has Opportunistic Infection, Has Dependent, Has contact person in case of emergency, & Availability of phone address for tracing

Others: - Weight, Initial ARV regimen provided, ART initiation duration, Total Months on ART and Months on ART where none adherence encountered

Operational definition of Non-Adherent :- if a client has at least one lost or dropped or restarted or stop or poor adherence record once he/she started taking ART therapy . Otherwise he/she will be considered as Adherent

3.6 Data Preparation

Anonymized data retrieved from Zewditu Memorial Hospital's ART database is collected under intake and follow-up categories. The Medical Record Number(MRN) is utilized as a unique identifier for each client record in every target dataset within the database. A comprehensive examination of the secondary data is conducted to establish familiarity with its content. The researcher actively identifies potential data quality issues, enabling the discovery of initial insights within the data.

In the data preparation phase, the following key steps are undertaken. First, the selection of data sets is determined, with careful documentation of the reasons for their inclusion or exclusion. Subsequently, the data cleaning process is initiated, which is often the most time-consuming task and crucial for preventing the "garbage-in, garbage-out" scenario. This step involves correcting, imputing, or removing erroneous values to enhance data quality. The construction of data follows, where new attributes are derived to provide additional insights. For instance, one might calculate body mass index by utilizing height and weight fields. Integration of data is the next step, involving the creation of new datasets through the combination of information from multiple sources. Lastly, formatting adjustments are made as needed, such as converting string values that store numbers into numeric values to facilitate mathematical operations. These comprehensive data preparation measures lay the foundation for effective analysis and meaningful insights (Ananda et al., 2022).

After gaining a comprehensive understanding of the available data and its limitations, the subsequent step involves preparing and analyzing the data to maximize its utility. The data preparation process is intricate and typically consumes approximately 80% of the project timeline. This involves describing data sources, performing statistical summary measures, identifying distinct values, and assessing the presence of missing values, outliers, and noisy data values. Decisions regarding the handling of missing values, outliers, and noisy data, as well as activities related to data transformation and reduction, are crucial aspects of this phase. Additionally, efforts are directed towards feature and attribute construction, involving the derivation of new attributes through segmentation, as emphasized in the work of Abdallah et al. (2017). This comprehensive approach to data preparation ensures the dataset's suitability for subsequent analysis and interpretation.

To meet the specifications of Weka 3.8.6, a series of select, update, and aggregate queries were executed to consolidate the database into a single table. This process aimed to extract the necessary attributes, ensuring that the data is appropriately formatted and ready for data mining operations. In consultation with domain expert 25 relevant fields(variables) from the data extracted from both Follow up card and Intake form electronic records selected with a total of 31,134 records. By excluding records of children less 15 years old and variables with majority null values, the total record reduced to 14,130 with 22 fields(variables) selected. Out of the 22 variables, three variables with around 50 % missing values are included as they are assumed to have significance for this study from previous similar research.

Table 1: Partial View of extracted data before cleaning

PatientGUI	PatientID	UANo	HIVConfirmDa	ARTStartDa	Sex	AgeAtDat	DateOfBir	BirthDa	BirthMon	BirthYe	BirthDateConfirme	PlaceOfBir
000009c72b094	364731	1.408E+11	10/1/2003	6/15/2004	M		00:00.0	NULL	NULL	NULL	NULL	
000a3b88b07d4	178597	1.409E+11	4/18/2013	8/16/2013	F	25	00:00.0	16	8	1988	NULL	
000d7dcc3fff46	35482	1.408E+11	6/18/2008	6/18/2008	F	10	00:00.0	25	11	1997	NULL	
002215f097374	364508	1.408E+11	8/13/2003	4/5/2004	M		00:00.0	NULL	NULL	NULL	NULL	
00258af005b54	362318	1.408E+11	12/14/2006	12/27/2006	F		00:00.0	NULL	NULL	NULL	NULL	
002951bf744f4	301819	14/09/020/0	5/15/2007	2/15/2008	F	38	00:00.0	10	3	1977	NULL	
002a34d0f10c4	534265	14/08/003/0	1/1/1900	9/29/2020	F	30	00:00.0	10	3	1990	NULL	
002a5c7571ec4	365180	1.408E+11	12/19/2004	12/19/2006	M		00:00.0	NULL	NULL	NULL	NULL	
00319338c2614	88444	1E+12	8/16/2011	8/6/2009	F		00:00.0	NULL	NULL	NULL	NULL	
0038e645128a4	30322	1.408E+11	8/11/2003	9/2/2003	M	36	00:00.0	2	9	1967	NULL	
0039f058fe264	488703	14/08/001/1	7/16/2019	7/16/2019	M	45	00:00.0	10	3	1974	NULL	
003a0d183e344	200139	1.709E+11	3/26/2012	5/25/2012	F		00:00.0	NULL	NULL	NULL	NULL	
003b0bac23764	44660	1.408E+11	4/7/2006	5/19/2006	M		00:00.0	NULL	NULL	NULL	NULL	
0045c6b780914	30359	1.408E+11	5/25/2007	5/30/2007	M	38	00:00.0	30	5	1969	NULL	
0045cae85f464	361896	1.408E+11	10/21/2007	11/9/2007	M		00:00.0	NULL	NULL	NULL	NULL	
004636eab5fb4	246060	1.408E+11	12/1/2014	12/4/2014	M	58	00:00.0	10	3	1956	NULL	
00470086b9ca4	363584	1.408E+11	6/13/2005	6/14/2005	F		00:00.0	NULL	NULL	NULL	NULL	

Table 2: Partial View of cleaned data

AgeatEnrolment	Sex	InitialWeight	InitialWHOstage	FunctionalStatus	ARTInitationSite	InitialRegimenNumber	ARTInitationDuration	MaritalStatus	EducationalLeve
41	F	40	4	Bedridden	TI	17	Beyond 2 Weeks	Never-Married	Primary
36	M	50	1	Working	InitiatedatHF	0	Beyond 2 Weeks	Married	Tertiary
32	F	45	3	Working	TI	0	Beyond 2 Weeks	Divorced	Secondary
40	F	70	3	Working	TI	0	Same Day	Married	Secondary
34	M	55	3	Working	TI	0	Beyond 2 Weeks	Married	Tertiary
27	M	45	3	Working	InitiatedatHF	0	Beyond 2 Weeks	Never-Married	Primary
49	F	45	4	Ambulatory	InitiatedatHF	0	Within 2 Weeks	Married	No education
24	F	40	3	Working	InitiatedatHF	0	Same Day	Married	Primary
32	M	55	2	Working	InitiatedatHF	0	Within 1 Week	Never-Married	Tertiary
52	M	55	1	Working	InitiatedatHF	0	Beyond 2 Weeks	Divorced	Tertiary
37	M	68	4	Working	TI	0	Same Day	Never-Married	Tertiary
30	F	64	1	Working	InitiatedatHF	0	Within 2 Weeks	Married	Secondary
25	F	45	3	Working	InitiatedatHF	0	Within 2 Weeks	Never-Married	Tertiary
40	M	60	4	Working	InitiatedatHF	0	Same Day	Married	Secondary
30	F	55	2	Working	InitiatedatHF	0	Beyond 2 Weeks	Married	Primary
32	F	52	2	Working	InitiatedatHF	0	Within 1 Week	Married	Secondary
32	F	43	2	Working	InitiatedatHF	0	Same Day	Married	Primary
35	F	66	2	Working	InitiatedatHF	0	Beyond 2 Weeks	Widowed	Primary

3.7 Data Analysis

3.7.1 Business Intelligence Tools

Business Intelligence refers to the use of technologies, processes, and tools to collect, analyze, and present business data in a meaningful and actionable manner. The primary goal of BI is to transform raw data into valuable insights, allowing organizations to make informed decisions, optimize performance, and gain a competitive advantage. BI encompasses a range of activities, including data integration, data cleansing, analysis, reporting, and visualization. By providing stakeholders with a consolidated view of key performance indicators, trends, and patterns, BI enables organizations to understand their business landscape, identify opportunities and challenges, and make strategic decisions that drive business success (Negash & Gray, 2008).

Business Intelligence tools are software applications designed to transform raw data into actionable insights, empowering organizations to make informed decisions. These tools gather, process, and analyze data from various sources, presenting it in a visual and easily understandable format. BI tools often include features such as reporting, dashboards, and data visualization, enabling users to explore trends, identify patterns, and gain a comprehensive understanding of their business performance. With capabilities for ad hoc querying, multidimensional analysis, and predictive analytics, BI tools cater to diverse user needs within an organization, from executives seeking strategic insights to operational staff monitoring day-to-day activities. By facilitating data-driven decision-making, BI tools play a crucial role in enhancing efficiency, uncovering opportunities, and fostering a culture of informed decision-making across all levels of an enterprise(Ranjan, 2009).

Even through there are several BI tools nowadays, the researcher selected Power BI as BI tools for this study. Power BI, developed by Microsoft, serves as a prominent business intelligence and analytics solution, enabling users to convert unprocessed data into valuable insights. Acknowledged for its intuitive interface and smooth compatibility with other Microsoft offerings, Power BI offers an extensive range of functionalities for data visualization, analysis, and reporting. Designed to accommodate users with diverse roles, from business analysts to data scientists, it facilitates connectivity to various data sources, the creation of interactive dashboards, and the seamless sharing of insights throughout the organization (Microsoft, 2023).

Power BI stands out as a versatile solution for organizations aiming to leverage their data effectively. It offers the flexibility to import data from various sources like databases, spreadsheets, and cloud services, allowing users to tailor and structure it according to their analytical requirements. With an intuitive drag-and-drop interface, users can effortlessly craft compelling visualizations, encompassing charts, graphs, and maps. The tool's real-time analytics feature ensures that users can track data changes as they occur, providing decision-makers with the most current information. Furthermore, the integration with Microsoft Azure enhances its capabilities, incorporating advanced analytics, machine learning, and artificial intelligence functionalities (Ferrari & Russo, 2016).

Power BI finds extensive application across various industries, aiding organizations in making data-driven decisions. In finance, it assists in financial reporting, budgeting, and forecasting. Healthcare professionals use Power BI to analyze patient outcomes, track medical trends, and enhance healthcare delivery. In sales and marketing, the tool helps optimize strategies by providing insights into customer behavior, sales performance, and market trends. The versatility and seamless integration with Microsoft's suite of products have positioned Power BI as a preferred choice for businesses aiming to derive actionable insights from their data.

3.7.2 Data Mining Modeling

Data mining involves the application of sophisticated algorithms and analytical techniques to sift through vast datasets, identifying hidden correlations and trends that may not be readily apparent through traditional methods. By leveraging the rich and diverse data housed in Zewditu Memorial hospital's ART database, data mining techniques can be employed to predict patient adherence to Antiretroviral Therapy. This process involves examining historical patient data, including treatment adherence records, demographic information, clinical outcomes, and other relevant variables. Through the application of predictive modeling and machine learning algorithms, data mining aims to create models that can forecast patient adherence patterns based on the identified data patterns. Ultimately, the insights derived from data mining in this context can contribute to the development of personalized intervention strategies, helping healthcare providers at Zewditu Memorial Hospital optimize Antiretroviral Therapy patient care, enhance treatment outcomes, and improve overall healthcare delivery.

3.7.2.1 Model Building

Model Building is the fourth phase of CRISP-DM. Here various models bult and assessed based on several different modeling techniques. The four tasks under this phase are **Select modeling techniques, Generate test design, Build model and Assess model** (Chapman et al. 2000):

The data modeling process involves several key steps. Firstly, one needs to decide on the modeling techniques, selecting algorithms like regression or neural nets that align with the objectives. Subsequently, a test design is generated, involving the division of data into training, test, and validation sets based on the chosen modeling approach. The next step is to build the model, which may involve executing code lines such as "reg = LinearRegression().fit(X, y)." Despite its seemingly glamorous nature, this step is pivotal in translating the chosen algorithm into a functional model. Finally, the data scientist assesses the model, considering multiple models in competition. Interpretation of results relies on domain knowledge, predefined success criteria, and the specifics of the test design.

Data preparation insights serve as the foundation for constructing diverse behavioral models. Employing a range of Machine Learning tools, multiple tests are conducted using this prepared data. The evaluation of each test's performance is then undertaken to gauge its effectiveness relative to other potential models.

Several blind tests are conducted to evaluate the model's performance against actual tests and its effectiveness with sample data. This iterative process often reveals opportunities for refining the model to enhance outcomes such as response rates, average purchase amount, customer reactivation, and more.

The collected data was first transferred to an Excel file (xlsx) and then converted to the Attribute-Relation File Format (arff) for readability in Weka version 3.9.0, a data mining software. Weka was chosen due to its availability as a free download online and the investigator's familiarity with the software. The J48 and Random Forest classifier algorithms within Weka were employed to train and test the predictive model. To enhance the model's efficiency, less significant attributes were eliminated before model fitting, and the resulting decision tree was presented for the selected model. Precision and recall measures were calculated and elucidated based on the confusion matrix.

3.7.2.2 Model Evaluation

The Evaluation phase of CRISP-DM looks more broadly at which model best meets the business and what to do next. The three tasks under this phase are Evaluate results, Review process, and Determine next steps (Chapman et al. 2000):

After evaluating the results of the modeling techniques and assessing their alignment with the business success criteria, the next step involves reviewing the overall process. This entails a thorough examination of the work undertaken, ensuring that no crucial steps were overlooked and that each phase was executed correctly. Findings from this review are summarized, and any necessary corrections are made. Subsequently, based on the insights gained and the success criteria met, the decision is made regarding the next steps. This may involve moving forward with deployment, refining the models through iteration, or initiating new projects to enhance the overall data-driven decision-making process.

The assessment of the dataset involves adjusting the percentage of training and testing data used for constructing the model. The classifiers' performance is then analyzed using various metrics, including different confusion matrices (True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNP), and False Negative Rate (FNR)). Additionally, the evaluation encompasses parameters such as the number of correctly classified instances, the number of leaves, the size of the trees, execution time, and the area under the ROC curve

CHAPTER FOUR

Results and Discussion

4.1Introduction

The final prepared data in excel is analyzed using Power BI for visualization and using Weka Data Mining tool for predictive model building. For Power BI data analysis default excel format (.xlsx) is used and for data mining model building the data is changed into .csv format and then converted to the Attribute-Relation File Format (arff) for readability in Weka version 3.9.0.

4.2 Data Visualization and Analysis Using Power BI

4.2.1 Statistical information of the Selected ART Clients Record

A review was conducted on records of 14,125 Antiretroviral Therapy (ART) clients out of the total 18,874 available within the specified period, resulting in a proportion of 74.8%. Of these, 13,858 clients (98.1%) were from Addis Ababa city, while the remaining 267 (1.9%) were from locations outside Addis Ababa. The average age of the respondents at the initiation of ART was 37.4 years, with a standard deviation of +9.65.

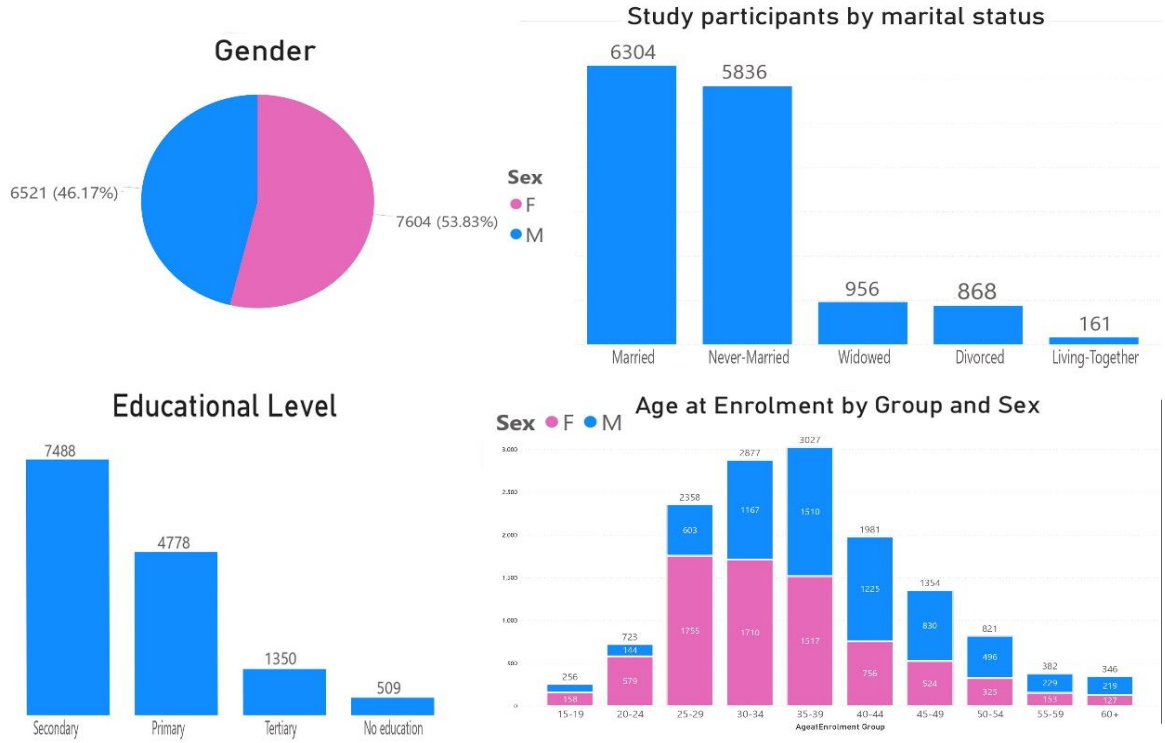

Figure 2: Statistical characteristics of variables Gender, Marital Status, Educational Level and Age group

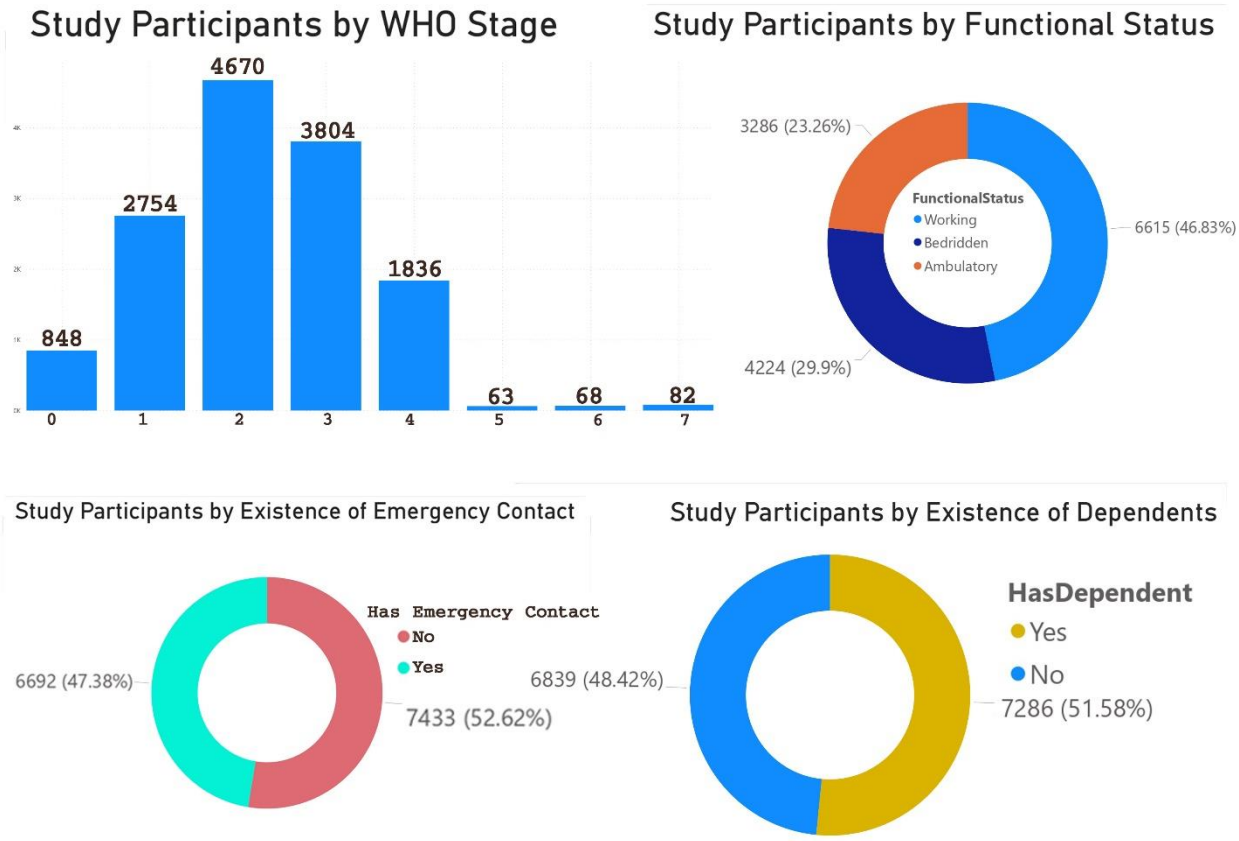

Figure 3: Statistical characteristics of variables WHO Stage, Functional Status, Emergency Contact and Dependents

Among the clients Orthodox religion followers accounts for 38.0%. There is a 7.6% higher proportion of females compared to males, while the number of married clients differs by 3.3% compared to those who have never been married. Opportunistic infections were present in 8.6% of clients at the initiation of Antiretroviral Therapy . Additionally, 47.4% of clients had an emergency contact person, and 51.6% of clients had dependents.

4.2.2 Adherence after Initiation to ART

Among the total study participants classified as adherent, 98% are located in Addis Ababa, while the remaining 2% reside outside Addis Ababa. The distribution is comparable for non-adherent clients, with an equal proportion from both Addis Ababa and locations outside the city. Regarding the presence of an emergency contact person, 50% of adherent clients had one, and an equal percentage did not. Likewise, for non-adherent clients, 45% had an emergency contact person, and 55% did not.

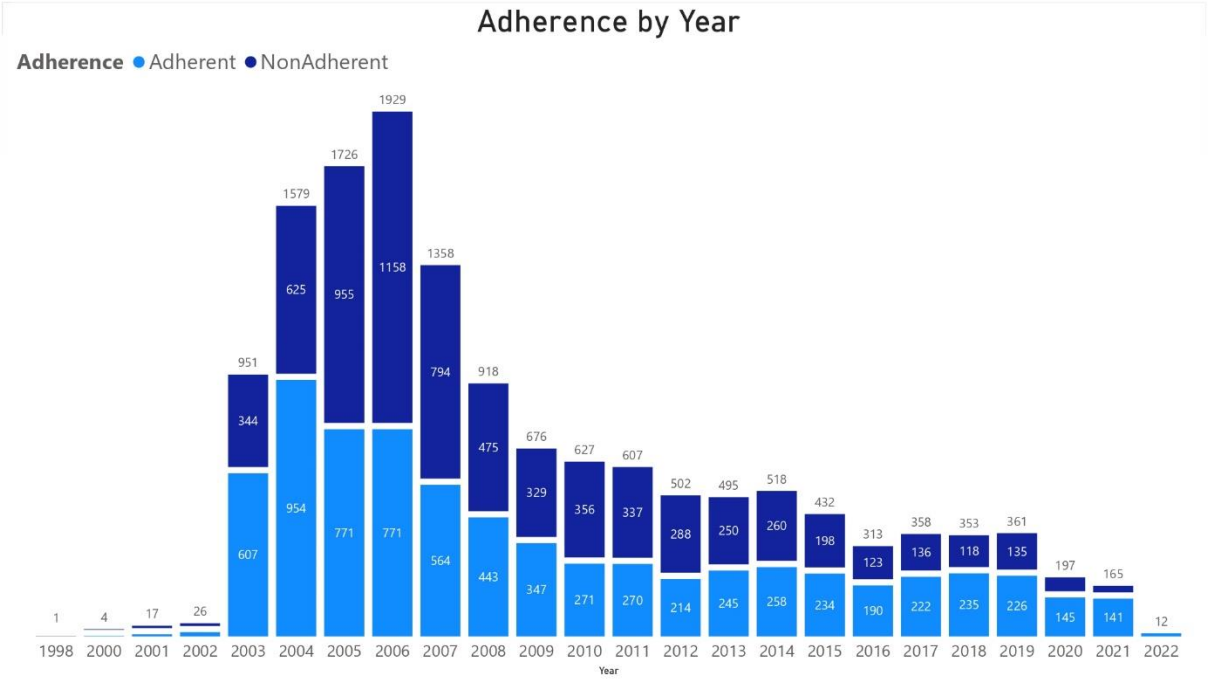

Figure 4: Adherence Status by ART Initiation Year

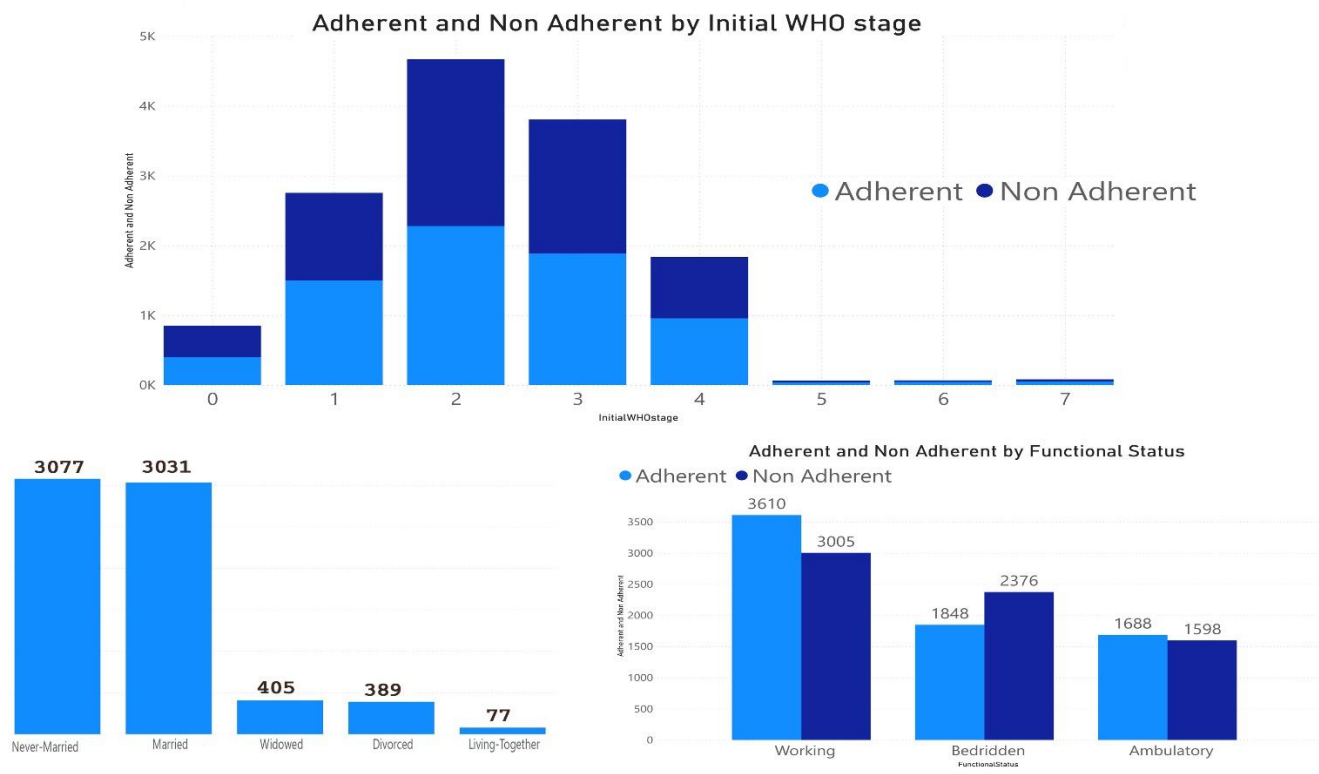

Figure 5: Adherence Status by WHO Stage, Marital and Working Status

4.2.3 ART Client Management Insight from the Power BI Visualization

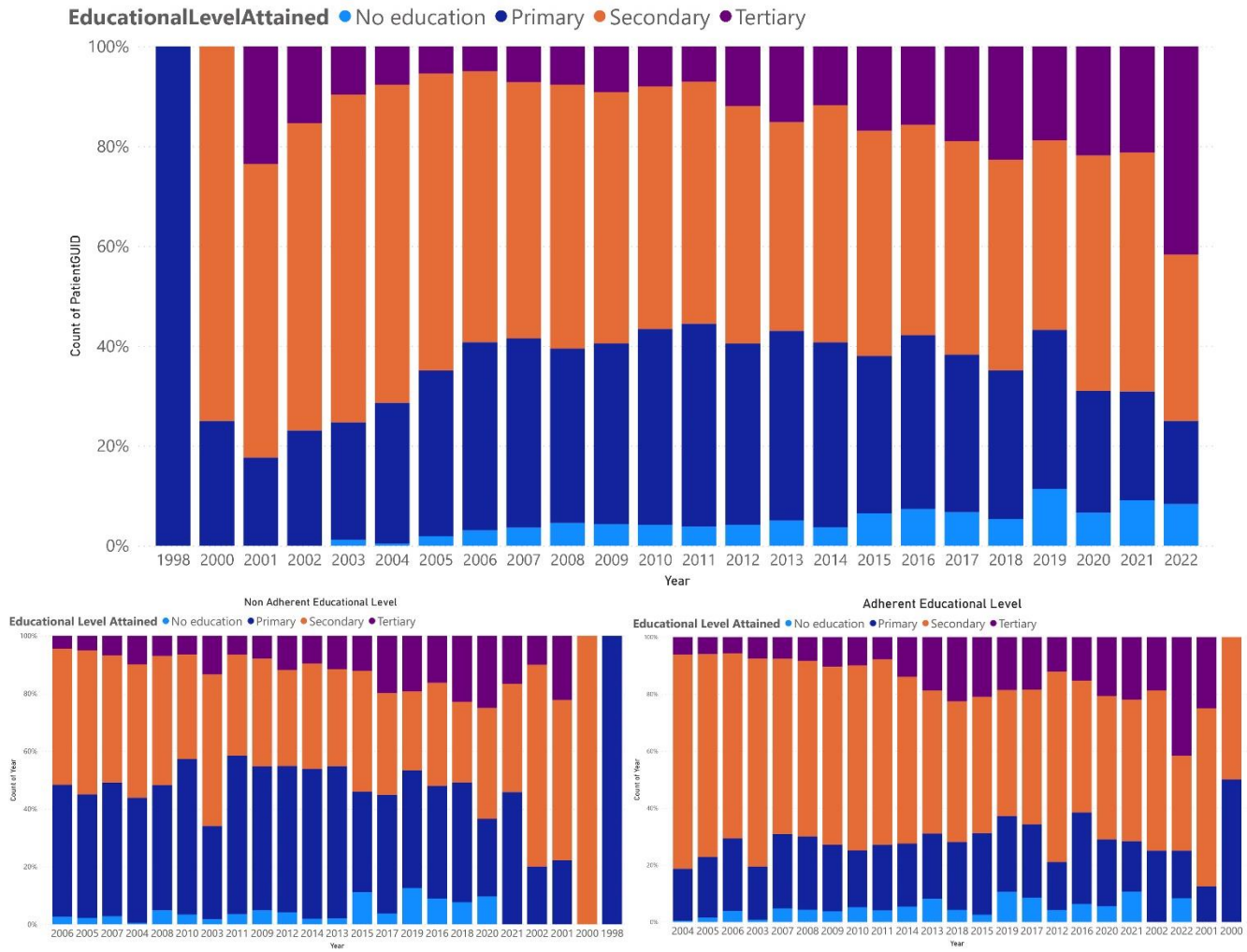

Figure 6: Trend of ART Initiation by Level of Education Attained and Adherence Status

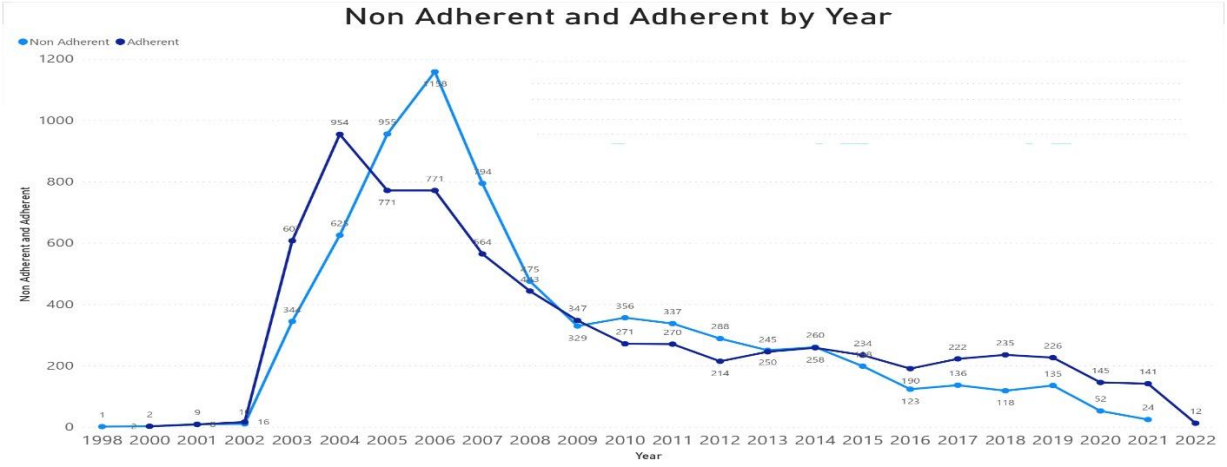

Figure 7: Trend of Adherence Status

As can be seen from figure 6 above, of the clients who initiated on ART in all ART initiation years majority of the clients attained primary and secondary level education. As the time become more recent clients who attained secondary level of education comply to ART treatment more than clients who attained primary level of education.

4.3 Data Mining Model Building

4.3.1 Data Mining Modeling Tests Using Weka 3.9.0

Utilizing Weka 3.9.0, multiple experiments were conducted employing the J48 and RandomForest algorithms. This section will delve into four experiments demonstrating superior performance. Two experiments involved constructing decision trees using the J48 algorithm, while the remaining two employed the RandomForest algorithm for comparison with J48. Table 3 and 4 below present the run parameters and outputs of the respective experiments, showcasing the use of two classification algorithms and two test modes in both tables.

One dataset consisting of 22 attributes used for all experiments. The total number of instances was 14,125 (50.59% Adherent and 49.41% non-Adherent). Two classifier algorisms; J4.8 and RandomForest were used in the experiments.

The two test modes conducted for both J48 and RandomForest:

a. Mod1: 10-fold cross-validation mode with all records
b. Mod2: Using 82% of the records to train the model and 18% of the record for testing the model.

Table 3: J48 Decision Tree experiments

Experiment	Input mode	Tree Size	Leaves count	Duration (sec.)	Accuracy (%)
1	Mod1	2540	1924	0.40	69.6
2	Mod2	2540	1924	0.41	71.6

After conducting a series of diverse experiments with J48, as illustrated in Table 3 above, it attains the highest classification accuracy of 71.6%.

Table 4: RandomForest experiments

Experiment	Input mode	Tree Size	Duration (sec.)	Accuracy (%)
3	Mod1	-	4.25	69.42
4	Mod2	-	0.3	69.75

The two additional experiments were conducted using the RandomForest algorithm, resulting in the best classification accuracy of 69.75%, as indicated in Table 4. Both algorithms achieved their highest accuracy when employing the Mod2 model-building approach, with the J48 decision tree algorithm surpassing the RandomForest classifier by 1.65%. Consequently, the model generated by J48 in the second experiment was chosen for further analysis..

In Experiment 2, the model was trained using the program's default values, except for the test options. A percent split was applied to divide the dataset into an 82% training set and an 18% test set, while other parameters under the "more parameter options" button were maintained at their default values. The results of the training are presented in the confusion matrix (Table 5) below..

Table 5: Confusion Matrix for J48 Algorithm with default parameter

	Predicted			Score
Actual	**Adherent**	**Non-Adherent**	**Total**	**(Accuracy rate)**
Adherent	907	350	**1257**	72.16%
Non-Adherent	372	913	**1285**	71.05%
Total	**1279**	**1263**	**2542**	**71.6%**

The confusion matrix, as illustrated in Table 5, indicates that out of the entire set of records processed by the program, 907 (72.16%) and 913 (71.05%) records were accurately classified in the Adherent and Non-Adherent classes, respectively. Conversely, 350 (27.84%) records were inaccurately classified as Non-Adherent when they should have been in the Adherent class, and 372 (28.95%) records were misclassified as Adherent when they actually belong to the Non-Adherent class. This indicates that out of the complete set of 2,542 test records, 1,820 (71.6%) records were accurately classified, while the remaining 722 (28.4%) records were classified incorrectly.

J48 pruned tree

```
------------------
EducationalLevelAttained = Secondary
|  AgeatEnrolment <= 40
|  |  MonthsonART <= 201
|  |  |  InitialRegimenCode = 1a: NonAdherent (1.0)
|  |  |  InitialRegimenCode = 1a30
|  |  |  |  Region = Addis Ababa
|  |  |  |  |  MaritalStatus = Never-Married
|  |  |  |  |  |  FunctionalStatus = Working
|  |  |  |  |  |  |  AgeatEnrolment <= 38
|  |  |  |  |  |  |  |  HasEmergencyContact = Yes
|  |  |  |  |  |  |  |  |  ARTInitationSite = InitiatedatHF
|  |  |  |  |  |  |  |  |  |  InitialWHOstage = 0: NonAdherent (0.0)
|  |  |  |  |  |  |  |  |  |  InitialWHOstage = 1: NonAdherent (5.0)
|  |  |  |  |  |  |  |  |  |  InitialWHOstage = 2
|  |  |  |  |  |  |  |  |  |  |  InitialWeight <= 48.5: Adherent (4.0/1.0)
……..
|  AgeatEnrolment > 40
|  |  AgeatEnrolment <= 50
|  |  |  HasEmergencyContact = Yes
|  |  |  |  Region = Addis Ababa
|  |  |  |  |  MonthsonART <= 122: Adherent (246.0/60.0)
|  |  |  |  |  MonthsonART > 122
|  |  |  |  |  |  InitialWeight <= 43: NonAdherent (16.0/2.0)
|  |  |  |  |  |  InitialWeight > 43
|  |  |  |  |  |  |  Religion = NULL
|  |  |  |  |  |  |  |  InitialRegimenCode = 1a: Adherent (0.0)
|  |  |  |  |  |  |  |  InitialRegimenCode = 1a30: Adherent (4.0/1.0)
|  |  |  |  |  |  |  |  InitialRegimenCode = 1a40: NonAdherent (4.0)
|  |  |  |  |  |  |  |  InitialRegimenCode = 1b: Adherent (0.0)
|  |  |  |  |  |  |  |  InitialRegimenCode = 1b30: NonAdherent (2.0/1.0)
|  |  |  |  |  |  |  |  InitialRegimenCode = 1b40: Adherent (1.0)
|  |  |  |  |  |  |  |  InitialRegimenCode = 1c
……
EducationalLevelAttained = No education
|  MonthsonART <= 27: Adherent (40.0/6.0)
|  MonthsonART > 27
|  |  InitialRegimenLine = 1
|  |  |  Occupation = NULL
|  |  |  |  InitialRegimenCode = 1a: Adherent (0.0)
|  |  |  |  InitialRegimenCode = 1a30
|  |  |  |  |  InitialWeight <= 47: NonAdherent (19.0/3.0)
|  |  |  |  |  InitialWeight > 47: Adherent (44.0/13.0)
|  |  |  |  InitialRegimenCode = 1a40: Adherent (3.0/1.0)
|  |  |  |  InitialRegimenCode = 1b: Adherent (0.0)
|  |  |  |  InitialRegimenCode = 1b30
|  |  |  |  |  PhoneAddressForTracingExist = Yes: NonAdherent (34.0/12.0)
|  |  |  |  |  PhoneAddressForTracingExist = No: Adherent (2.0)
|  |  |  |  InitialRegimenCode = 1b40: NonAdherent (2.0/1.0)
|  |  |  |  InitialRegimenCode = 1c
|  |  |  |  |  Sex = M
|  |  |  |  |  |  ARTInitationDuration = Beyond 2 Weeks: Adherent (7.0)
|  |  |
```

Figure 8: Partial View of J48 Decision Tree Model with Default Parameter

While this training approach has demonstrated commendable accuracy, the decision tree appears intricate and extensive, as depicted in Figure 8. The tree's size and the number of leaves generated in this training are 2540 and 1924, respectively. Consequently, extracting comprehensible rules requires navigating through numerous nodes of the tree.

Thus, in order to streamline the generation of rule sets for experiment 2 and enhance clarity, the researcher endeavored to adjust the default parameter values of the selected classifier to reduce the size of the tree and the number of leaves. With this aim, the parameter minNumObj (minimum number of instances in a leaf) was modified to 10, deviating from its default value of 2, and this value was determined after several trial adjustments. Consequently, this adjustment implies that the classification process continues until the number of records at each leaf reaches 10.

The outcomes of employing the J48 algorithm with an 82% - 18% training-test instance split are encapsulated and outlined in Table 6 below. The findings of this experiment reveal that, as a result of tweaking certain parameters, the tree's size has diminished to 871, and the number of leaves has decreased to 714.

Table 6:Confusion Matrix for J48 Algorithm with Modifying Default Parameter Values

	Predicted			Score
Actual	**Adherent**	**Non-Adherent**	**Total**	**(Accuracy rate)**
Adherent	931	326	**1257**	74.06%
Non-Adherent	400	885	**1285**	68.87%
Total	**1331**	**1211**	**2542**	**71.44.%**

Among the 2,542 test data fed into the program, 1,816 (71.44%) records were accurately classified, while the remaining 726 (28.56%) records were misclassified. The outcomes of this training approach appear to facilitate the generation of simpler rules with a comparable accuracy rate to the model constructed prior to modifying the default parameter values.

Hence, the model created in experiment 2, employing an 82%-18% training-test instance split mode, utilizing the J48 algorithm, and with the parameter minNumObj (minimum number of instances in a leaf) set to 10 (from its default value of 2), is chosen as the ultimate operational model. A partial view of the associated decision tree is depicted in Figure 9 below, and the complete tree is appended in Annex V.

The decision tree model, as indicated in Figure 9 below, has been chosen, and it is crafted with 871 nodes and 714 leaves, incorporating twenty-one optimal attributes. The figures enclosed in brackets at the end of the tree (leaf node) denote the count of records correctly and incorrectly classified in each class.

Of the twenty-one optimal attributes, the model employs the clients' level of education attained (EducationalLevelAttained) as the primary classification criterion for determining the adherence status of a client. The second-best classification criterion chosen by the model, depending on the level of education attained, is the age of the clients when initiated to ART (AgeatEnrolment). To meet additional classification requirements, the remaining nineteen attributes are utilized in constructing the model at different levels as classification criteria to ascertain the adherence status of a client.

J48 pruned tree

```
------------------
EducationalLevelAttained = Secondary
|  AgeatEnrolment <= 40
|  |  MonthsonART <= 201
|  |  |  InitialRegimenCode = 1a: NonAdherent (1.0)
|  |  |  InitialRegimenCode = 1a30
|  |  |  |  FunctionalStatus = Working
|  |  |  |  |  InitialWHOstage = 0: Adherent (1.0)
|  |  |  |  |  InitialWHOstage = 1: NonAdherent (16.0/8.0)
|  |  |  |  |  InitialWHOstage = 2: NonAdherent (59.0/26.0)
|  |  |  |  |  InitialWHOstage = 3
|  |  |  |  |  |  HasEmergencyContact = Yes: NonAdherent (68.0/26.0)
|  |  |  |  |  |  HasEmergencyContact = No: Adherent (28.0/7.0)
|  |  |  |  |  InitialWHOstage = 4: NonAdherent (21.0/6.0)
|  |  |  |  |  InitialWHOstage = 5: NonAdherent (0.0)
|  |  |  |  |  InitialWHOstage = 6: NonAdherent (0.0)
|  |  |  |  |  InitialWHOstage = 7: NonAdherent (0.0)
|  |  |  |  FunctionalStatus = Bedridden
|  |  |  |  |  PhoneAddressForTracingExist = Yes: NonAdherent (100.0/30.0)
|  |  |  |  |  PhoneAddressForTracingExist = No
|  |  |  |  |  |  InitialWHOstage = 0: NonAdherent (6.0/2.0)
|  |  |  |  |  |  InitialWHOstage = 1: NonAdherent (30.0/12.0)
|  |  |  |  |  |  InitialWHOstage = 2
|  |  |  |  |  |  |  MonthsonART <= 194: NonAdherent (89.0/32.0)
|  |  |  |  |  |  |  MonthsonART > 194
|  |  |  |  |  |  |  |  InitialWeight <= 47: NonAdherent (13.0/5.0)
|  |  |  |  |  |  |  |  InitialWeight > 47: Adherent (18.0/5.0)
|  |  |  |  |  |  InitialWHOstage = 3
|  |  |  |  |  |  |  ARTInitationDuration = Beyond 2 Weeks
|  |  |  |  |  |  |  |  InitialWeight <= 45: Adherent (24.0/8.0)
|  |  |  |  |  |  |  |  InitialWeight > 45: NonAdherent (31.0/9.0)
|  |  |  |  |  |  |  ARTInitationDuration = Same Day: Adherent (19.0/6.0)
|  |  |  |  |  |  |  ARTInitationDuration = Within 1 Week: NonAdherent (14.0/6.0)
|  |  |  |  |  |  InitialWHOstage = 4: NonAdherent (2.0/1.0)
|  |  |  |  |  |  InitialWHOstage = 5: NonAdherent (0.0)
|  |  |  |  |  |  InitialWHOstage = 6: NonAdherent (0.0)
|  |  |  |  |  |  InitialWHOstage = 7: NonAdherent (0.0)
|  |  |  |  FunctionalStatus = Ambulatory: NonAdherent (283.0/104.0)
|  |  |  InitialRegimenCode = 1a40
|  |  |  |  Religion = NULL: NonAdherent (57.0/22.0)
|  |  |  |  Religion = Protestant: Adherent (3.0)
```

```
| | | | Religion = Orthodox: Adherent (17.0/7.0)
| | | | Religion = Muslim: NonAdherent (2.0)
| | | | Religion = Other: NonAdherent (0.0)
| | | | Religion = Catholic: NonAdherent (0.0)
| | | | Religion = PROTESTANT: NonAdherent (0.0)
| | | | Religion = ORTHODOX: NonAdherent (0.0)
| | | | Religion = MUSLIM: NonAdherent (0.0)
| | | | Religion = CATHOLIC: NonAdherent (0.0)
| | | InitialRegimenCode = 1b: Adherent (1.0)
| | | InitialRegimenCode = 1b30
| | | | AgeatEnrolment <= 24: NonAdherent (57.0/15.0)
| | | | AgeatEnrolment > 24
| | | | | Religion = NULL
| | | | | | HasEmergencyContact = Yes: Adherent (12.0/2.0)
| | | | | | HasEmergencyContact = No
| | | | | | | FunctionalStatus = Working: Adherent (19.0/9.0)
| | | | | | | FunctionalStatus = Bedridden
| | | | | | | | ARTInitationDuration = Beyond 2 Weeks
| | | | | | | | | InitialWeight <= 38: NonAdherent (11.0/2.0)
| | | | | | | | | InitialWeight > 38: Adherent (115.0/41.0)
| | | | | | | | ARTInitationDuration = Same Day
| | | | | | | | | MonthsonART <= 181: Adherent (14.0/3.0)
| | | | | | | | | MonthsonART > 181: NonAdherent (30.0/14.0)
| | | | | | | | ARTInitationDuration = Within 1 Week: NonAdherent (28.0/12.0)
```

Figure 9: Partial View of J48 Decision Tree with Default Parameter Values Modified

Table 7: Precision and Recall Accuracy Measures for J48 Algorithm with Default Parameter Values Modified

Class values	Precision	Recall
Adherent	0.699	0.741
Non-Adherent	0.731	0.689
Average	**0.715**	**0.714**

The selected model undergoes testing with 2,542 test instances, constituting 18% of the total instances, to evaluate its predictive performance. The resultant precision and recall measures are outlined in Table 7 above.

Precision and recall serve as widely utilized classification metrics in the field of data mining. In the context of a classification task, precision for a class is determined by the ratio of true positives (i.e., items accurately labeled as part of the positive class) to the total number of elements labeled as part of the positive class (i.e., the sum of true positives and false positives, representing items incorrectly labeled as part of the class). Conversely, recall is defined as the ratio of true positives to the total number of elements that genuinely belong to the positive class (i.e., the sum of true positives and false negatives, which denote items not labeled as part of the positive class but should have been).

Hence, the chosen model in this investigation produces precision measures of 69.9% and 73.1% for the Adherent and Non-Adherent class-labels, respectively (refer to Table 6 and Table 7). This signifies that among the instances classified by the model as Adherent, 69.9% (931 out of 1,331 instances) genuinely belong to the Adherent class. Similarly, among the instances classified by the model as Non-Adherent, 73.1% (885 out of 1,211 instances) truly pertain to the Non-Adherent class.

The classification performance of the model, expressed through the recall measure for the Adherent and Non-Adherent classes, is 74.1% and 68.9%, respectively. This indicates that concerning instances truly belonging to the Adherent class, the model correctly classified them as Adherent 74.1% of the time. Likewise, with respect to instances genuinely belonging to the Non-Adherent class, the model accurately classified them as Non-Adherent 68.9% of the time..

4.3.2 ART Client Management Insight from the J48 Decision Tree

The insights acquired from the decision tree created in experiment 2, utilizing adjusted parameters, can be articulated as a series of rules. Obtaining these rules involves traversing the decision tree, generating a rule for each leaf, and amalgamating all the tests identified along the path from the root to the leaf node. The subsequent rules are a subset of those extracted from the decision tree illustrated in Figure 9.

1. IF EducationalLevelAttained = Secondary AND AgeatEnrolment <=40 AND MonthsonART <= 201 AND InitialRegimenCode = 2i

THEN NonAdherent (4.0/2.0)

From this rule, we can deduce that clients undergoing the initial regimen with 2i code regimen, having a duration of ART treatment less than or equal to 201 months on ART, and an age during ART initiation less than or equal to 40 years, coupled with a secondary level of education, exhibit non-adherence or noncompliance to the treatment 67% of the time. This noncompliance may stem from the potential side effects of the prescribed medicine. Consequently, clients possessing these specific characteristics necessitate special attention from healthcare service providers.

2. IF EducationalLevelAttained = Secondary AND AgeatEnrolment <=40 AND MonthsonART <= 201 AND InitialRegimenCode = 1j

THEN Adherent (152.0/31.0)

From this rule, we observe that clients who share similar characteristics as outlined in rule 1, except for the ARV regimen they are taking (1j), tend to comply with the treatment 83% of the time. Consequently, clients with these specific characteristics may not necessitate special attention from service providers.

3. IF EducationalLevelAttained = No education AND MonthsonART <= 27

THEN Adherent (40.0/6.0)

From this rule, we understand that clients with no education at all, along with an ART treatment duration of less than or equal to 27 months, adhere or comply with the treatment 87% of the time. Consequently, clients with these characteristics may not necessitate special attention from service providers.

4. IF EducationalLevelAttained = Tertiary AND Region = Oromia AND MonthsonART <= 97:

THEN Adherent (19.0/4.0)

From this rule, we learn that clients with a tertiary level of education, originating from the Oromia region, and with an ART treatment duration less than or equal to 97 months, adhere or comply with the treatment 82.6% of the time. Consequently, clients possessing these characteristics may not necessitate special attention from service providers.

5. IF EducationalLevelAttained = Tertiary AND Region = Oromia AND MonthsonART > 97

THEN NonAdherent (20.0/6.0)

In this rule, we note that clients with attributes similar to those specified in rule 5, except for MonthsonART being greater than 97 months, demonstrate non-compliance with the treatment 78% of the time. Therefore, clients possessing these particular characteristics may not require special attention from service providers.

6. IF EducationalLevelAttained = Tertiary AND Region = Amhara

THEN Adherent (8.0)

In this rule, we note that clients with a tertiary level of education and originating from Amhara region comply with the treatment 100% of the time. Therefore, clients possessing these particular characteristics may not require special attention from service providers.

7. IF EducationalLevelAttained = Tertiary AND Region = Tigray

THEN NonAdherent (3.0/1.0)

In this rule, we note that clients with a tertiary level of education and originating from Tigray region do not comply with the treatment 75% of the time. Thus, clients with these characteristics require especial attention by service providers.

8. IF EducationalLevelAttained = PRIMARY

THEN Nonadherent (169.0/66.0)

From this rule we learn that clients with primary level of education are noncompliant to the treatment 72% of the time. Thus, clients with these characteristics require especial attention by service providers.

9. IF EducationalLevelAttained = Tertiary AND Region = Addis Ababa AND HasDependent = No AND AgeatEnrolment <= 28 years

THEN NonAdherent (57.0/6.0)

From this rule, we learn that clients with a tertiary level of education, originally from Addis Ababa city, have no dependent and with an age during ART initiation less or equal to 28 years do not adhere or comply with the treatment 90.5% of the time. Thus, clients with these characteristics require especial attention by service providers and find a mechanist helps them to adhere to the treatment.

The figures enclosed in parentheses indicate the count of records correctly and incorrectly classified within each class. For instance, in rule 2, the numbers within the parentheses signify that this rule accurately classified 152 records as Adherent and inaccurately classified 31 records as Adherent when they actually belonged to the Non-Adherent category.

The above rules represent a portion of the insights derived from the chosen model for predicting clients' adherence status. For instance, according to rule 8, if a client possesses a primary level of education, their adherence to therapy is unlikely without considering other predictor attributes.

Therefore, special attention is needed to either maintain their medication regimen or enhance their compliance with the therapy. Similarly, in rule 7, clients with a tertiary level of education originating from the Tigray region are predicted to exhibit low adherence to therapy under the current service provision strategy at the hospital, thus warranting special attention. Additionally, in accordance with the recommendations of domain experts, the algorithm-generated rule suggesting that clients residing outside the health facility catchment area are less likely to adhere to treatment aligns with the domain experts' observations.

4.4 ART Client Management Insight from Power BI and Decision Tree Combined

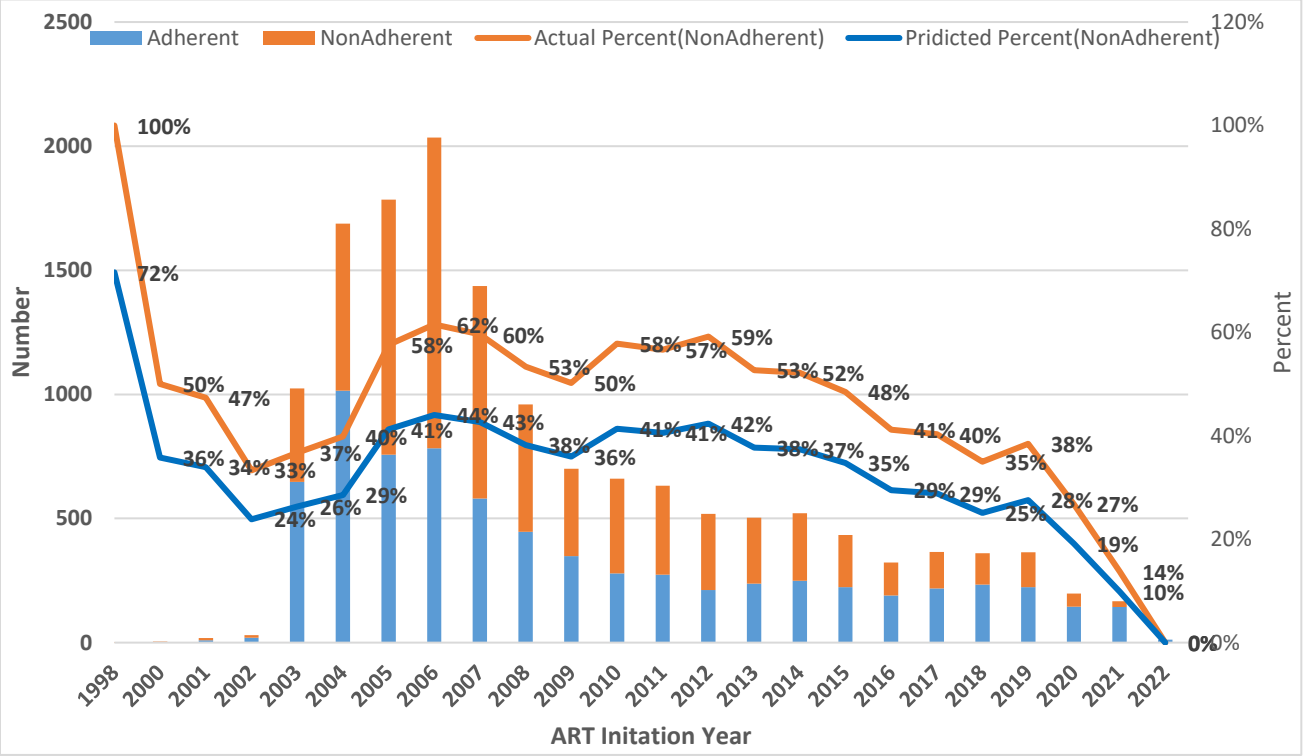

Figure 10: Power BI Visualization and Data Mining Predication Combined

As depicted on figure 10 above by combining trends from power BI and data mining modeling prediction, we can improve the treatment outcomes of 71.44 % of the clients predicated to be non-adherent per each ART initiation year by taking proactive measures including customized counseling and social supports.

4.5 Result Comparison with Similar Researches

To contextualize the findings of this study within existing research, we can draw parallels with a study titled "Determinants of adherence to antiretroviral therapy among HIV-positive adults in sub-Saharan Africa" by Heestermans et al. This study highlights the interconnectedness of sociodemographic, psychosocial, health status, treatment-related, and intervention-related factors in fostering optimal adherence to ART, aligning closely with the findings of this research. However, it is noteworthy that while Heestermans et al. conducted a systematic literature review to analyze determinants of adherence, this study employs Data Mining techniques based on actual experimental data. This distinction underscores the empirical nature of this research, offering insights gleaned directly from data analysis, which may complement and enrich their findings.

In a study titled "Incidence and Predictors of ART Failure among Adults Receiving HIV Care at Zewditu Memorial Hospital, Addis Ababa, Ethiopia" by Chalachew et al., predictors for ART failure were identified, including WHO stage 3 at the start, baseline CD4 count, disclosure, and opportunistic infections. These predictors partially overlap with those identified in this research. However, it is notable that Chalachew et al. utilized the SPSS statistical package, which has limited predictive capabilities compared to Data Mining techniques employed in this study. Furthermore, their research was based on data collected from a smaller sample of 595 participants, whereas this study utilized electronic data from the ART client database encompassing over 14,000 clients. This distinction in sample size and analytical approach underscores the robustness and scalability of this findings, offering insights derived from a larger and more comprehensive dataset, thus potentially enhancing the generalizability and applicability of the results.

In another study titled "Adherence to Antiretroviral Treatment and Associated Factors among Seropositive People Received Treatment in Jimma Town Public Health Facilities, Ethiopia" by Zenash et al., the researchers investigated factors associated with adherence to treatment among 385 survey participants. Their findings highlighted various factors such as food security, substance abuse, use of other medication, relationship with healthcare providers, and irregular appointments as determinants of treatment adherence. However, it is worth noting that the sample size in their study was relatively small, and they utilized the SPSS tool, which has limited capacity compared to the Data Mining techniques employed in our research. Despite these differences, both studies

contribute valuable insights into the factors influencing adherence to antiretroviral treatment in Ethiopia, albeit with variations in methodology and scope.

4.6 Benefits and challenges of implementing Business Intelligence

Benefits of Implementing BI in ART Client Management at Zewditu Memorial Hospital

Enhanced Decision-Making: Implementing Business Intelligence in Antiretroviral Therapy client management at Zewditu Memorial Hospital offers the potential for more informed and timely decision-making. Through advanced data visualization and analytics, healthcare providers can gain valuable insights into patient outcomes, treatment trends, and resource utilization. This can lead to more effective strategies in optimizing patient care, ensuring medication adherence, and addressing specific healthcare challenges.

Optimized Resource Allocation: Business Intelligence tools provide the capability to analyze resource utilization patterns and identify areas for optimization. In a resource-constrained environment like Zewditu Memorial Hospital, where efficient allocation is critical, Business Intelligence can assist in identifying cost-effective interventions, reducing wastage, and ensuring that limited resources are directed toward the most impactful areas of ART client management. This can lead to improved operational efficiency and a maximized impact on patient outcomes.

Improved Program Accountability: With the donor-driven nature of ART services in Ethiopia, Business Intelligence implementation can contribute to improved program accountability. By aligning data collection and reporting with program requirements, the hospital can not only meet donor expectations but also demonstrate the impact of the funds invested. Business Intelligence can aid in tracking key performance indicators, adherence to program guidelines, and overall program effectiveness, fostering a transparent and accountable healthcare environment.

Challenges of Implementing BI in ART Client Management at Zewditu Memorial Hospital

Resource Constraints: Zewditu Memorial Hospital operates in a resource-constrained setting, and the adoption of Business Intelligence technologies may face challenges related to infrastructure, expertise, and financial resources. The initial investment required for Business Intelligence implementation, including training staff and acquiring suitable technology, may strain the hospital's limited resources.

Poor Data Quality and Completeness: The availability of poor-quality data, characterized by inaccuracies and incompleteness, poses a significant challenge. In an environment where data capture is primarily driven by reporting requirements and lacks awareness of data-driven decision-making, ensuring the reliability and completeness of the data becomes crucial. Business Intelligence tools are only as effective as the data they analyze, making data quality improvement a prerequisite for successful implementation.

Limited Data-Driven Decision-Making Awareness: The awareness and understanding of data-driven decision-making principles may be low among healthcare professionals at Zewditu Memorial Hospital. Overcoming this challenge involves not only implementing Business Intelligence tools but also fostering a culture of data-driven decision-making. Adequate training and awareness programs are essential to ensure that healthcare providers can leverage Business Intelligence tools effectively to drive positive outcomes in Antiretroviral Therapy client management.

Donor-Driven Reporting vs. Comprehensive Data Capture: The existing data capture system in Antiretroviral Therapy services is primarily geared towards meeting reporting requirements set by donors and programs. This may result in a lack of comprehensive data capture, missing out on particular patient insights that could contribute to improved healthcare outcomes. Aligning donor-driven reporting with the broader goals of comprehensive data capture for informed decision-making presents a unique challenge in this context.

Addressing these challenges requires a strategic and phased approach to Business Intelligence implementation, considering the specific context of Zewditu Memorial Hospital. It involves building both technological and human capacity while navigating through complex donor-driven healthcare environment.

CHAPTER FIVE

Conclusion and Recommendation

5.1 Conclusion

The analysis of the current state of ART client management at Zewditu Memorial Hospital underscores a notable underutilization of its vast Electronic Medical Record(EMR) data for ART clients. Despite accumulating a substantial amount of data over several years, the hospital predominantly uses this information mainly for producing program reports to meet donor demands. However, this narrow focus fails to capitalize on the potential of data-driven decision-making to elevate patient care outcomes. Bridging this gap necessitates a comprehensive evaluation of how the hospital can effectively leverage its EMR data repository to inform and optimize decision-making processes, thereby enhancing patient care.

In pursuit of the second research objective, an exploration and selection process was undertaken to identify BI tools that best suited for managing ART client data and related healthcare processes. After careful consideration, Power BI emerged as the optimal choice for visualization purposes, offering robust capabilities to dissect and present complex healthcare data in a comprehensible manner. Additionally, Data Mining Decision Tree were found to be instrumental in unearthing non-trivial, hidden decision support information from extensive datasets collected over time.

In evaluating how BI applications can aid in identifying treatment trends, adherence patterns, and patient outcomes for ART clients, this study underscores the pivotal role of BI tools like Power BI, offering visual summaries that simplify the identification of patterns and trends in healthcare data, providing valuable insights. Additionally, Data Mining Decision Tree was selected to predict ART client adherence to treatment, leveraging its advanced algorithms to uncover patterns and trends within the data. This proactive approach enables tailored counseling and additional social support to promote patient adherence to ART treatment which is vital for treatment success.

In analyzing the long-term benefits and challenges of implementing BI in ART client management at Zewditu Memorial Hospital, several key findings emerged. The implementation of BI holds promise for enhanced decision-making processes, providing healthcare providers with valuable insights into patient outcomes, treatment trends, and resource utilization. This can result in more effective strategies for optimizing patient care and ensuring medication adherence. Moreover, BI tools facilitate optimized resource allocation, crucial in a resource-constrained environment like

Zewditu Memorial Hospital, where efficient allocation is essential for maximizing impact on patient outcomes. However, the implementation of BI also presents challenges, including resource constraints such as infrastructure, expertise, and financial resources. Poor data quality and completeness, coupled with limited awareness of data-driven decision-making principles among healthcare professionals, further compound these challenges. Additionally, the existing donor-driven reporting system may hinder comprehensive data capture necessary for informed decision-making. Overcoming these challenges necessitates a strategic and phased approach to BI implementation, involving the development of both technological and human capacity within the complex donor-driven healthcare environment. Thus, while BI offers significant potential for improving ART client management, addressing these challenges is paramount for successful long-term implementation and realization of its benefits at Zewditu Memorial Hospital.

While the study effectively addressed its research questions, it also illuminated significant new inquiries regarding the transformative potential of electronic health record data in enhancing healthcare services in Ethiopia through modern information technology. As the healthcare landscape evolves, leveraging modern information technology, particularly through BI techniques and data mining, becomes increasingly vital for optimizing patient care, resource allocation, and operational efficiency.

5.2 Recommendations

The findings of this research, while contributing to academic knowledge, hold practical implications for Antiretroviral Therapy care providers. The insights gained can aid care providers in identifying high-risk targets within the Antiretroviral Therapy care system and addressing issues that may impede the continuity of Antiretroviral Therapy care. This application of data-driven decision-making can significantly enhance the healthcare management outcomes, cost-effective resource allocation, and enhanced operational efficiency.

To enhance the scope and applicability of future research, data from more Heath Facilities need to be included. There are currently over 800 Health Facilities where EMR-ART application implemented through Ethiopia. Even though this study is conducted for academic purpose, it is envisioned as a cornerstone for future research endeavors in the field of ART datasets. Future studies could expand their reach to include a more comprehensive dataset from various healthcare facilities, thereby enriching the knowledge base and contributing to more robust insights.

Stakeholders, including health institutions, regional health bureaus, the Ministry of Health, and local and international partners involved in funding and providing technical assistance for HIV pandemic control, should collaboratively work to improve data completeness and quality for electronic medical recording systems. The current focus on generating reports at Zewditu Hospital and similar government health institutions in Ethiopia should evolve towards an emphasis on enhancing the overall data completeness and quality of Electronic Medical Record systems like Enhanced EMR-ART. This shift would unlock valuable information from extensive patient records, either archived or live, fostering a culture of data-driven decision-making for more effective HIV control strategies.

5.3 Future research direction

This study explored utilization of data-driven decision-making for Antiretroviral Therapy Client Management at Zewditu Memorial Hospital, utilizing the available electronic client records from the HIV patient database. To expand the depth of understanding, the researcher suggests future studies that encompass additional Antiretroviral Therapy service-providing health facilities across the county. Such an approach would facilitate the creation of more insightful visualizations and predictive models, that can significantly impact Antiretroviral Therapy service provision on a broader scale. Furthermore, enhancing the research could involve introducing a qualitative component, thus ensuring a more comprehensive exploration of the subject matter.

References

World Health Organization (WHO). Everybody's business - strengthening health systems to improve health outcomes: WHO's framework for action. WHO; Geneva: 2007. http://www.who.int/healthsystems/strategy/ everybodys_business.pdf

Guidelines for Management of Opportunistic Infections and Anti Retroviral Treatment in Adolescents and Adults in Ethiopia. Addis Ababa: FHAPCO and FMOH; 2007.

UNAIDS 2030 Ending the AIDS epidemic

ICAP-Ethiopia SI department bi annual report April 2021on Enhanced EMR-ART Application coverage.

Zewditu Memorial Hospital Wikipedia: https://en.wikipedia.org/wiki/Zewditu_Hospital

Abdallah, Z. S., Du, L., & Webb, G. I. (2017).Data Preparation.

Abou-Stait,F. (2005). Working Paper 76-Are exports the Engine of Economic Growth? An application of Cointegration and causality Analysis for Egypt,1977-2003 (No.211).

Acito, F., &Khatri, V. (2014). Business analytics: Why now and what next?.

Akhtar, N., Perwej, A., &Perwej, Y. (2020).Data analytics and visualization using Tableau utilitarian for COVID-19 (Coronavirus). Global Journal of Engineering and Technology Advances, 3(2), 028-050.

Al-Hajj, S., Fortuno III, E. S., & Fisher, B. (2011). Data Visualization of Immunological Competence of HIV

Exposed but Uninfected (HEU) infants. IEEE Visual Analytics for Science and Technology (VAST).

Alharthi, N., &Gutub, A. (2017).Data visualization to explore improving decision-making within Hajj services. Scientific Modelling and Research, 2(1), 9-18.

Amabile, S., Laghzaoui, S., Peignot, J., Peneranda, A., &Boudrandi, S. (2013). Business intelligence practices for exporting SMEs. International Business Research, 6(2), 101.

Single Point HIV Prevalence Estimate [Internet]. Addis Ababa: Federal Ministry of Health/National HIV/AIDS Prevention and Control Office; [cited 2011 Mar 27]. Available from: http://www.etharc.org/aidsineth/publications/singlepointprev_2007.pdf

Ethiopian Public Health Institute. HIV Related Estimates and Projection in Ethiopia. 2019; (3-4)

Tiyou A, Belachew T, Alemseged F, Biadgilign S. Predictors of adherence to antiretroviral therapy among people living with HIV/AIDS in resource limited setting of southwest ethiopia. AIDS Research and Therapy. 2010;7:39

Hecht FM, Grant RM, Petropoulos CJ, Dillon B, Chesney MA, Tian H. Sexual transmission of an HIV-1 variant resistant to multiple Reverse transcriptase and protease inhibitors. N Engl J Med. 1998;339(5):307-11.

Endrias M, Alemayehu W, Gail D. Adherence to ART in PLWHA at Yigalem Hospital South Ethiopia. Ethiop.J.Health Dev. 2008;22 (2):174-9.

Zewditu Memorial Hospital ART clinic; fact sheet and monthly report. 2018;(March).

Fayyad U, Piatetsky-Shapiro G, Smyth P. From Data Mining to Knowledge Discovery in Databases. AI MAGAZINE. 1996;:37 - 54.

Han J, Kamber M. Data mining: concepts and techniques. 2nd ed. Los Altos, CA: Morgan Kaufman; 2006.

Hian Chye Koh and Gerald Tan, "Data Mining Applications in Healthcare", Journal of Healthcare Information Management — Vol.19,No.2

Milley, A. (2000), "Healthcare and Data mining", Health Management Technology,21(8),44-47

Biafore, S. (1999), "Predictive solutions bring more power to decision makers", Health Management Technology,20(10),12-14

Ciosa K J, Mooree GW. Uniqueness of medical data mining. Artificial Intelligence in Medicine. 2002;26:1–24.

Hand DJ. Data Mining: Statistics and More? The American Statistician. 1998 May;52(2):112 - 8.

Silver M., Sakata T., Hua-Ching SU., Hua-Ching C., Dolins SB., Dolins MJ. Case Study: How to Apply Data Mining Techniques in a Healthcare Data Warehouse. Journal of Healthcare Information Management. 2001;15(2):155-64.

Lin RS, Rhee SY, Shafer RW, Das AK. A combined data mining approach for infrequent events:

analyzing HIV mutation changes based on treatment history. Comput Syst Bioinformatics Conf. 2006;:385-8.

Prosperi MCF, Rosen-Zvi M, Altmann A, Giambenedetto SD, Kaiser R. Antiretroviral Therapy Optimisation without Genotype Resistance Testing: A Perspective on Treatment History Based Models. PLoS ONE. 2010 Oct;5(10):1-8.

Singh Y, Mars M. Support vector machines to forecast changes in CD4 count of HIV-1 positive patients. Scientific Research and Essays. 2010 Sep;5(17):2384-2390.

Seara T. Support vector machine prediction of HIV-1 drug resistance using the viral nucleotide patterns. Transactions of the Royal Society of South Africa. 2009;64(1):62 - 72.

Federal Democratic Republic of Ethiopia Ministry of Health (FMoH). 2010. Health Sector Development Program IV, 2010/11–2014/14. FMOH Addis Ababa; Ethiopia http://www.nationalplanningcycles.org/sites/default/files/ country_docs/Ethiopia/ethiopia_hsdp_iv_final_draft_2010_-2015.pdf

Pronovost et al. 2015, Creating a high-reliability health care system: improving performance on core processes of care at Johns Hopkins Medicine. Acad Med 2015 Feb;90(2):165-72.

Chapman et al. 2000. CRISP-DM 1.0

Lankowski, Alexander J., et al. "Impact of geographic and transportation-related barriers on HIV outcomes in sub-Saharan Africa: a systematic review." AIDS and Behavior 18.7 (2014): 1199-1223.1182-1190.

Martin, et al, 2005 “The challenge of patient adherence”

Wirth, R. & Hipp, Jochen. (2000). CRISP-DM: Towards a standard process model for data mining. Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining.

R. Sharma and P. Srinath, "Business Intelligence using Machine Learning and Data Mining techniques - An analysis," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2018, pp. 1473-1478, doi: 10.1109/ICECA.2018.8474847

Muraina, Ishola & Ahmad, Azizah. (2012). Healthcare Business Intelligence: The Case of

University's Health Center.

Abdul-Kareem, S.Et. Al. 2000. ANN as a Tool for Medical Prognosis.

Kaur, Harleen, and Siri Krishan Wasan. "Empirical study on applications of data mining techniques in healthcare." Journal of Computer science 2.2 (2006): 194-200.

Fayyad, Usama, Gregory Piatetsky-Shapiro, and Padhraic Smyth. "From data mining to knowledge discovery in databases." AI magazine 17.3 (1996): 37-37.

Chesney, M.A. (2000) Factors Affecting Adherence to Antiretroviral Therapy. Clinical Infectious Diseases, 30, S171-S176. https://doi.org/10.1086/313849

https://azure.microsoft.com/en-us/overview/what-are-business-intelligence-tools/

https://mopinion.com/business-intelligence-bi-tools-overview/

https://online.hbs.edu/blog/post/importance-of-business-analytics.

https://research.aimultiple.com/data-cleaning/

https://searchbusinessanalytics.techtarget.com/definition/data-preparation

Annexes

Annex I: Intake Form-A

HIV Care/ ART clinic Intake form A Social demographics & family care FEDERAL MINISTRY OF HEALTH OF ETHIOPIA

MRN 040231 Enrollment Date 01/02/2004 UANo 03080121542 Unit TB No

Facility Beletshachew Health Center

Client Name DESSIA TILAHUN NEGUSSIE Date Of Birth 02/02/1971 Sex M

Mothers Name Place of Delivery

Mode of Delivery Marital Status M

Religion Orthodox

Level of Education Occupation

Address

CLIENT REFERAL INFORMATION Outside the health facility Health Center

CARE GIVER/ EMERGENCY CONTACT INFORMATION

Full Name Age Sex Relation

DISCLOSURE

Does anyone else know about your/your child's HIV status?

If yes who knows your/your child's HIV +ve status?

FAMILY MEMBERS HIV STATUS

Does the client has family members (spouse / parent and child / siling < 15 years)
If yes, Check HIV status at ENROLLMENT and on FOLLOW UP visit as er table below

At enrollment				**At enrollment or follow up visit**: update all black fields as status changed						
Family Member	Age	Sex	Health status (healthy, Chronically ill, dead)	Counseled for HIV*	Tested for HIV*	HIV test Result: NR, R *	Enrolled in Care*	Started ART	Unique ART*	Remark
Spouse/Parent										
Child/Sibiling 1										
Child/Sibiling 2										
Child/Sibiling 3										
Child/Sibiling 4										
Child/Sibiling 5										

Annex II: Intake Form-B

HIV Care / ART clinic Inake form - B Past & presenting illness FEDERAL MINISTRY OF HEALTH ETHIOPIA

MRN 040231 FacilityName Beletshachew Health Center

PAST OPPORTUNISTIC ILLNESS (OI), PROPHYLAXIS AND TREATMENT

Past OI Fever (>1 month and unexplained)

Past Prophylaxis / Treatment

Cotrimoxazole: Date started: If stopped, date stopped

INH: Date If stopped, date stopped

ART/TI: If yes, Regimen Start Date Months on ART

Still on Treatment CotriStartDateET

Client STATUS AT ENROLLMENT

VITAL SIGNS Temp (C) PULSE/HR(/m) BP (mmHg) RR(R/m)

ANTROPOMETRY Height (cm) Weight (Kg) BMI (kg/m2) Head Circf.

Functional Status: Working **Developmental status (for a child < 5)**

PRESENTING SYMPTOM

CLIENT'S PREGNANCY STATUS at Enrollment Pregnant? LMP EDD

CLIENT GENERAL APPEARANCE OF AT Enrollment

System	Status	Specify Abnormal Finding
HEENT	Normal	
Lymph nodes	Normal	
Chest	Normal	
Heart	Normal	
Abdomen	Normal	
Genitourinary System	Normal	
Musculo-skeletal system	Normal	
Skin	Normal	
Nervous System	Normal	

WHO HIV Clinical Stage at enrollment: WHO Stage 3 why:

ART ELIGIBILITY ASSESSMENT AND PLAN

Eligible for ART if yes why

Annex III: Follow-up Form

HIV CARE/ART FOLLOW-UP FORM **FEDERAL MINISTRY OF HEALTH**

Facility Name_____

Name: _____ Age _____years (Months for Children <5 years) Sex ☐ M F ☐

Address: Region_____ Woreda/Sub city_____ Kebele_____ House No._____Telephone _____

Patient Card No_____ Unique ART No. _____ Date confirmed HIV+ (Retesting): _____/____/_____(DD/MM/YY)

Type of HIV Test ☐ Rapid HIV tests ☐ DNA/PCR (for children) Client readiness: (date client is ready)_____/_____/_____(DD/MM/YY) Height (Adult) in cm: _____

S/ US	Follow up date dd/mm/yy	Months on ART	Weight (Kg) Edema+/-	BMI (MUAC for pregnant woman or bedridden)	Pregnancy status /FP method	Functional Status W,A,B	WHO Stage (1-4 / T1- T4	TB Screen - P/N	Xpert MTB/RIF Sent (Y)	Xpert MTB/RIF Result P/N	TB Treatment / TB Prophylaxis	Cervical Cancer Screened (VIA) (0,1,2,3,4)	OIs/ HIV related cancers	Assessed & Managed for pain	CD4 /mm^3 or % if < 5 Yrs.	VL result (copies/ ml)	Hgb. ALT/ AST/ Cr/ CrAg	Co-trimoxazole Preventive therapy Adherence (G,F,P)	Co-trimoxazole Preventive therapy Dispense dose	Fluconazole preventive therapy(FPT) (Dispensed Dose)	Other medications / or nutritional supplements dispensed	Assess & counsel for ARV Adherence Adherence (G,F,P)	Assess & counsel for ARV Adherence Why Fair or Poor?	ARV drugs Dispense Dose/ code	ARV drugs Side effect	ARV drugs Reason for change	Client sets HIV prevention plan Disclosure for Children	Next Visit Date dd/mm/yy
					If child Length/ Height/ Head Circumference for < 3years	Developmental Milestone (A.D.R)																						

Version: Hamle_ 2011

Annex IV: Follow-up Form-Back Side

S/US

S=Scheduled US=Unscheduled

TB SCREEN

SCREEN FOR TB AT EVERY VISIT

Adult & Adolescent

1. Current Cough?
2. Fever?
3. Night sweats.
4. Weight loss?

P= (Positive screen)-Yes to any of the above---Evaluate for TB.

N= (Negative screen)- No to all the questions above---assess for IPT eligibility

Children 0-14 years old

1. Current Cough?
2. Fever?
3. Weight loss or poor weight gain ?
4. Contact history with TB patient?

P=(Positive screen)-Yes to any one of the four---evaluate for TB

N=(Negative screen)-No to all four---assess for eligibility to IPT

Xpert MTB/ RIF (GeneXpert)

P= Positive

N= Negative

TB PROPHYLAXIS TREATMENT

INH1-6=Currently on INH prophylaxis (Number refers to months on INH)
INH-C=Completed treatment
INH- DC=discontinued for any reason
TB RX 1-6 =Currently on AntiTB (numbers refer to months on Rx)
TB RX DC=discontinued TB Rx for any reason
TB Rx C=Completed anti TB Rx

OI/ Opportunistic Cancers

NOI= No OI or Opportunistic cancer
Z=Zoster
BP=Bacterial Pneumonia
PTB= Pulmonary Tuberculosis
EPTB= Extra pulmonary tuberculosis
TO= Thrush oral
EC= esophageal candidiasis
UM=ulcers-mouth
DC or DA=Diarrhea Chronic/Acute
PCP=Pneumocystis pneumonia
CT= CNS Toxoplasmosis
CM=Cryptococcal Meningitis
NHL=NonHodgkins Lymphoma
KS=Kaposi's Sarcoma
CCa=Cervical cancer
O=Other

Months on ART

Duration in months since initiation of ART:
0 = ART Initiation date
1 week = 1 week
2 weeks = 2 weeks
3 weeks = 3 weeks
1 = 1 month
2= 2 months
If not started on ART (Pre-ART) leave this column blank

Pain Assessment & Management

Assess for Pain & Manage as
NP= No pain
S1= WHO Step 1
S2= WHO step 2
S3= WHO step 3

Nutritional Status (adults)

BMI (wt/(ht²)) (for non-pregnant / non postpartum)
1= Normal (18.5-24.99kg/m²)
2= Mild (17-18.49kg/m²)
3=Moderate malnutrition (16 -16.99kg/m²)
4=Severe malnutrition (< 16kg/m²)
5=Over weight (25-29.99 kg/m²)
NB: write the codes (1,2,3,4 or 5)

MUAC
(for pregnant/ postpartum /bedridden)
1= Normal (>23cm)
2=Moderate malnutrition (19-23cm)
3=Severe malnutrition (<19 cm for pregnant and postpartum /<18cm for bedridden)

ADHERENCE

Estimate adherence using the table below:

Adheren	%	# missed doses	
G (Good)	≥95%	(of 30 doses) <2 doses	(of 60 doses) ≤3 doses
F (Fair)	85-94%	2-4 doses	4-9 doses
P (Poor)	<85%	≥5 doses	≥10 doses

If Fair or Poor adherence, in why column note reason:
1. Toxicity/ Side effects
2. Share with others
3. Forgot
4. Felt better
5. Too ill
6. Stigma, discloser
7. Drug stock out
8. Lost/ ran out of pills
9. Delivery/ travel problems
10. Inability to pay
11. Alcohol
12. Depression
13. Other

Pregnancy Status /Family Planning Method

P = Pregnant (If pregnant, give estimated due date (EDD))
PMTCT = Referred to PMTCT & indicate linkage
WP = want to become pregnant
No FP = not pregnant & is not using any FP methods
FP= On Family Planning (enter code):
1= Condoms
2= Oral contraceptive pills
3= Injectable
4= Implant
5=Intrauterine device
6=Vasectomy/ tubal legation
7= Abstinence (no sex)

Length / height/ HC

Measure length / height in cm for children younger than 14 years at EVERY visit.
Measure head circumference in cm for children younger than 3 years of age at EVERY visit

Nutritional Status (Children)

W/H
1=Normal/ Appropriate (> -1 Z- score)
2=Mild (< -1 and > -2 Z- score)
3=MAM - Moderate Acute Malnutrition (< -2 and > -3 Z-score)
4=SAM-Severe Acute Malnutrition (< -3Z-score)

SIDE EFFECTS

1. No side Effects
2. Nausea
3. Diarrhea
4. Fatigue
5. Headache
6. Numbness/ tingling/pain
7. Rash
8. Anemia
9. Abdominal pain
10. Jaundice
11. Fat changes
12. Dizzy, anxiety, nightmare, depression
13. Other

REASONS FOR REGIMEN CHANGE

1. Toxicity/ Side effects
2. Due to new TB
3. New drug available
4. Drug stock out
5. Clinical failure
6. Immunologic failure
7. Virologic failure
8. Other

Functional status

W=Working (able to perform usual work in or out of the house, harvest, go to school or, for children, normal activities or playing)
A=Ambulatory (able to perform activities of daily Living)
B=Bedridden (not able to perform activities of daily living)

DEVELOPMENTAL MILESTONES FOR CHILD

A= Appropriate:
Sitting without support .....3 to 9months
Standing with assistance.....5 to 11months
Hands and knees crawling 6 to1 3months
Walking with assistance .....7 to 14 months
Standing alone.....8 to 17 months
Walking alone..... 9 to 18 months

Delay: Failure to attain milestones for age

Regression: Loss of what has been attained for age

Nutritional Status (Older children & adolescents

BMI for age(5-18yrs)
1=Normal/appropriate (> -1 Z score)
2=Mild (< -1 and > -2 Z score)
3=Moderate malnutrition (< -2 and > -3 Z score)
4=Severe malnutrition (< -3 Z score)
*BMI for age for older children and adolescents.

REASONS FOR STOPPING REGIMEN

STOP = Stopped ART
If STOP, In why column, note reason:
1. Toxicity/side effects
2. Treatment failure
3. Poor adherence
4. Illness, hospitalization
5. Drugs out of stock
6. Patient lack finances
7. Other patient decision
8. Other

VIRAL LOAD

Mark '*' under VL column when requested/ specimen collected,
Write the amount and interpret as undetectable and detectable for clinical intervention

Cervical Cancer Screening

Screened with VIA (1- 4):
0. VIA Not Done
1. VIA Negative
2. VIA Positive: eligible for Cryo
3. VIA Positive: Non-Eligible for Cryo
4. Suspicious for Cervical Cancer

Client readiness

Enter the date (dd/mm/yy) client is ready for ART initiation when client is counseled, adherence barriers addressed and client is willing to start ART

Client Set HIV Prevention Plan
D = Agreed to Disclose to partner/ family / friend,
PT= Agreed to bring partner for testing
ChT= agreed to bring children for testing,
SSex= discussed & agreed to practice safer sex
SubU=Decides to avoid or decrease Substance use
ASS= Assessed for STI
SRX=client managed for STI

For children Fill in stage of HIV disclosure
DS0 = No disclosure
DS1=Satge1, about the illness, taking medicine, keeping healthy,
DS2=Stage2, about germs, body soldiers,
DS3=Stage3, use of terms like CD4, Viral Load, HIV

DISPENSE DOSE/ REGIMEN CODE

ADULT FIRST LINE	CHILD FIRST LINE
1c = AZT-3TC-NVP 1d = AZT - 3TC - EFV 1e = TDF - 3TC - EFV 1f = TDF + 3TC + NVP 1g = ABC + 3TC + EFV 1h = ABC + 3TC +NVP 1j = TDF + 3TC + DTG 1K = AZT + 3TC + DTG 1i = Other specify	4c = AZT+ 3TC+NVP 4d = AZT+3TC+EFV 4e = TDF+3TC+EFV 4f = AZT +3TC + LPV/r 4g = ABC + 3TC + LPV/r 4i = TDF + 3TC + DTG 4j = ABC + 3TC + DTG 4K = AZT + 3TC + DTG 4L = ABC + 3TC + EFV 4h = Other specify
ADULT SECOND LINE	**CHILD SECOND LINE**
2e= AZT +3TC +LPV/r 2f =AZT+3TC +ATV/r 2g=TDF + 3TC+-LPV/r 2h= TDF + 3TC + ATV/r 2i= ABC + 3TC+ LPV/r 2j= Other specify	5e=ABC+3TC+LPV/r 5f=AZT + 3TC + LPV/r 5g=TDF + 3TC + EFV 5h=ABC + 3TC + EFV 5i= TDF + 3TC+LPV/r 5k= RAL+AZT+3TC 5L= RAL+ABC+3TC 5m= ABC+3TC+DTG 5n= AZT+3TC+DTG 5j= Other specify
ADULT THIRD LINE	**CHILD THIRD LINE**
3a = DRV/r+DTG+AZT+3TC 3b = DRV/r+DTG+TDF+3TC 3d = DRV/r+ABC+EFV+3TC 3c = Other specify	6a = DRV/r + RAL + AZT +3TC 6b = DRV/r + RAL + TDF +3TC 6c = DRV/r + DTG + AZT +3TC 6d = DRV/r + DTG +TDF+3TC 6F = DRV/r+DTG+ABC+3TC 6e = Other specify

In the follow-up date, in 2nd column if one of the options below applies, use raw next to the last visit to enter the appropriate information:
TO=Transfer out LOST=not seen since ≥1 month,3 months DROP= lost to follow-up for >3 months STOP= When the clinician stop ART for different reason and patient is on follow up DEAD

Annex V: Data Mining Decision Tree for selected Model

J48 pruned tree

```
------------------
EducationalLevelAttained = Secondary
|  AgeatEnrolment <= 40
|  |  MonthsonART <= 201
|  |  |  InitialRegimenCode = 1a: NonAdherent (1.0)
|  |  |  InitialRegimenCode = 1a30
|  |  |  |  FunctionalStatus = Working
|  |  |  |  |  InitialWHOstage = 0: Adherent (1.0)
|  |  |  |  |  InitialWHOstage = 1: NonAdherent (16.0/8.0)
|  |  |  |  |  InitialWHOstage = 2: NonAdherent (59.0/26.0)
|  |  |  |  |  InitialWHOstage = 3
|  |  |  |  |  |  HasEmergencyContact = Yes: NonAdherent (68.0/26.0)
|  |  |  |  |  |  HasEmergencyContact = No: Adherent (28.0/7.0)
|  |  |  |  |  InitialWHOstage = 4: NonAdherent (21.0/6.0)
|  |  |  |  |  InitialWHOstage = 5: NonAdherent (0.0)
|  |  |  |  |  InitialWHOstage = 6: NonAdherent (0.0)
|  |  |  |  |  InitialWHOstage = 7: NonAdherent (0.0)
|  |  |  |  FunctionalStatus = Bedridden
|  |  |  |  |  PhoneAddressForTracingExist = Yes: NonAdherent (100.0/30.0)
|  |  |  |  |  PhoneAddressForTracingExist = No
|  |  |  |  |  |  InitialWHOstage = 0: NonAdherent (6.0/2.0)
|  |  |  |  |  |  InitialWHOstage = 1: NonAdherent (30.0/12.0)
|  |  |  |  |  |  InitialWHOstage = 2
|  |  |  |  |  |  |  MonthsonART <= 194: NonAdherent (89.0/32.0)
|  |  |  |  |  |  |  MonthsonART > 194
|  |  |  |  |  |  |  |  InitialWeight <= 47: NonAdherent (13.0/5.0)
|  |  |  |  |  |  |  |  InitialWeight > 47: Adherent (18.0/5.0)
|  |  |  |  |  |  InitialWHOstage = 3
|  |  |  |  |  |  |  ARTInitationDuration = Beyond 2 Weeks
|  |  |  |  |  |  |  |  InitialWeight <= 45: Adherent (24.0/8.0)
|  |  |  |  |  |  |  |  InitialWeight > 45: NonAdherent (31.0/9.0)
|  |  |  |  |  |  |  ARTInitationDuration = Same Day: Adherent (19.0/6.0)
|  |  |  |  |  |  |  ARTInitationDuration = Within 1 Week: NonAdherent (14.0/6.0)
|  |  |  |  |  |  InitialWHOstage = 4: NonAdherent (2.0/1.0)
|  |  |  |  |  |  InitialWHOstage = 5: NonAdherent (0.0)
|  |  |  |  |  |  InitialWHOstage = 6: NonAdherent (0.0)
|  |  |  |  |  |  InitialWHOstage = 7: NonAdherent (0.0)
|  |  |  |  FunctionalStatus = Ambulatory: NonAdherent (283.0/104.0)
|  |  |  InitialRegimenCode = 1a40
|  |  |  |  Religion = NULL: NonAdherent (57.0/22.0)
|  |  |  |  Religion = Protestant: Adherent (3.0)
|  |  |  |  Religion = Orthodox: Adherent (17.0/7.0)
|  |  |  |  Religion = Muslim: NonAdherent (2.0)
|  |  |  |  Religion = Other: NonAdherent (0.0)
|  |  |  |  Religion = Catholic: NonAdherent (0.0)
|  |  |  |  Religion = PROTESTANT: NonAdherent (0.0)
|  |  |  |  Religion = ORTHODOX: NonAdherent (0.0)
|  |  |  |  Religion = MUSLIM: NonAdherent (0.0)
|  |  |  |  Religion = CATHOLIC: NonAdherent (0.0)
|  |  |  InitialRegimenCode = 1b: Adherent (1.0)
|  |  |  InitialRegimenCode = 1b30
|  |  |  |  AgeatEnrolment <= 24: NonAdherent (57.0/15.0)
|  |  |  |  AgeatEnrolment > 24
|  |  |  |  |  Religion = NULL
|  |  |  |  |  |  HasEmergencyContact = Yes: Adherent (12.0/2.0)
|  |  |  |  |  |  HasEmergencyContact = No
|  |  |  |  |  |  |  FunctionalStatus = Working: Adherent (19.0/9.0)
|  |  |  |  |  |  |  FunctionalStatus = Bedridden
|  |  |  |  |  |  |  |  ARTInitationDuration = Beyond 2 Weeks
```

```
|   |   |   |   |   |   |   |   |   InitialWeight <= 38: NonAdherent (11.0/2.0)
|   |   |   |   |   |   |   |   |   InitialWeight > 38: Adherent (115.0/41.0)
|   |   |   |   |   |   |   |   ARTInitationDuration = Same Day
|   |   |   |   |   |   |   |   |   MonthsonART <= 181: Adherent (14.0/3.0)
|   |   |   |   |   |   |   |   |   MonthsonART > 181: NonAdherent (30.0/14.0)
|   |   |   |   |   |   |   |   ARTInitationDuration = Within 1 Week: NonAdherent (28.0/12.0)
|   |   |   |   |   |   |   FunctionalStatus = Ambulatory
|   |   |   |   |   |   |   |   InitialWHOstage = 0: NonAdherent (1.0)
|   |   |   |   |   |   |   |   InitialWHOstage = 1: Adherent (7.0/2.0)
|   |   |   |   |   |   |   |   InitialWHOstage = 2
|   |   |   |   |   |   |   |   |   PhoneAddressForTracingExist = Yes: NonAdherent (18.0/6.0)
|   |   |   |   |   |   |   |   |   PhoneAddressForTracingExist = No
|   |   |   |   |   |   |   |   |   |   InitialWeight <= 41: Adherent (19.0/5.0)
|   |   |   |   |   |   |   |   |   |   InitialWeight > 41
|   |   |   |   |   |   |   |   |   |   |   InitialWeight <= 51.5: NonAdherent (40.0/11.0)
|   |   |   |   |   |   |   |   |   |   |   InitialWeight > 51.5: Adherent (16.0/5.0)
|   |   |   |   |   |   |   |   InitialWHOstage = 3
|   |   |   |   |   |   |   |   |   MaritalStatus = Never-Married
|   |   |   |   |   |   |   |   |   |   MonthsonART <= 189: NonAdherent (16.0/3.0)
|   |   |   |   |   |   |   |   |   |   MonthsonART > 189: Adherent (14.0/6.0)
|   |   |   |   |   |   |   |   |   MaritalStatus = Divorced: NonAdherent (0.0)
|   |   |   |   |   |   |   |   |   MaritalStatus = Married: Adherent (25.0/11.0)
|   |   |   |   |   |   |   |   |   MaritalStatus = Widowed: NonAdherent (0.0)
|   |   |   |   |   |   |   |   |   MaritalStatus = Living-Together: NonAdherent (0.0)
|   |   |   |   |   |   |   |   InitialWHOstage = 4: NonAdherent (1.0)
|   |   |   |   |   |   |   |   InitialWHOstage = 5: NonAdherent (0.0)
|   |   |   |   |   |   |   |   InitialWHOstage = 6: NonAdherent (0.0)
|   |   |   |   |   |   |   |   InitialWHOstage = 7: NonAdherent (0.0)
|   |   |   |   |   Religion = Protestant: NonAdherent (11.0/5.0)
|   |   |   |   |   Religion = Orthodox
|   |   |   |   |   |   InitialWeight <= 57.5: NonAdherent (87.0/30.0)
|   |   |   |   |   |   InitialWeight > 57.5: Adherent (17.0/4.0)
|   |   |   |   |   Religion = Muslim: NonAdherent (2.0/1.0)
|   |   |   |   |   Religion = Other: Adherent (0.0)
|   |   |   |   |   Religion = Catholic: NonAdherent (1.0)
|   |   |   |   |   Religion = PROTESTANT: Adherent (0.0)
|   |   |   |   |   Religion = ORTHODOX: Adherent (0.0)
|   |   |   |   |   Religion = MUSLIM: Adherent (0.0)
|   |   |   |   |   Religion = CATHOLIC: Adherent (0.0)
|   |   |   InitialRegimenCode = 1b40: Adherent (26.0/3.0)
|   |   |   InitialRegimenCode = 1c
|   |   |   |   HasDependent = Yes
|   |   |   |   |   Religion = NULL: Adherent (64.0/19.0)
|   |   |   |   |   Religion = Protestant
|   |   |   |   |   |   MaritalStatus = Never-Married: Adherent (11.0/3.0)
|   |   |   |   |   |   MaritalStatus = Divorced: NonAdherent (4.0/2.0)
|   |   |   |   |   |   MaritalStatus = Married: NonAdherent (22.0/5.0)
|   |   |   |   |   |   MaritalStatus = Widowed: NonAdherent (4.0/2.0)
|   |   |   |   |   |   MaritalStatus = Living-Together: NonAdherent (0.0)
|   |   |   |   |   Religion = Orthodox
|   |   |   |   |   |   InitialWHOstage = 0: Adherent (10.0/2.0)
|   |   |   |   |   |   InitialWHOstage = 1
|   |   |   |   |   |   |   MonthsonART <= 167: Adherent (32.0/11.0)
|   |   |   |   |   |   |   MonthsonART > 167: NonAdherent (13.0/3.0)
|   |   |   |   |   |   InitialWHOstage = 2
|   |   |   |   |   |   |   HasPastOI = Yes: Adherent (16.0/4.0)
|   |   |   |   |   |   |   HasPastOI = No
|   |   |   |   |   |   |   |   AgeatEnrolment <= 36: NonAdherent (57.0/26.0)
|   |   |   |   |   |   |   |   AgeatEnrolment > 36: Adherent (12.0/1.0)
|   |   |   |   |   |   InitialWHOstage = 3
|   |   |   |   |   |   |   HasPastOI = Yes: NonAdherent (20.0/4.0)
|   |   |   |   |   |   |   HasPastOI = No
```

```
| | | | | | | | Sex = M: NonAdherent (18.0/8.0)
| | | | | | | | Sex = F: Adherent (34.0/13.0)
| | | | | | InitialWHOstage = 4: NonAdherent (35.0/15.0)
| | | | | | InitialWHOstage = 5: Adherent (1.0)
| | | | | Religion = Muslim: NonAdherent (14.0/3.0)
| | | | | Religion = Other: Adherent (1.0)
| | | | | Religion = Catholic: Adherent (0.0)
| | | | | Religion = PROTESTANT: Adherent (0.0)
| | | | | Religion = ORTHODOX: Adherent (0.0)
| | | | | Religion = MUSLIM: Adherent (0.0)
| | | | | Religion = CATHOLIC: Adherent (0.0)
| | | | HasDependent = No
| | | | | AgeatEnrolment <= 21: NonAdherent (15.0/6.0)
| | | | | AgeatEnrolment > 21: Adherent (285.0)
| | | InitialRegimenCode = 1d
| | | | HasDependent = Yes
| | | | | HasEmergencyContact = Yes
| | | | | | InitialWHOstage = 0: NonAdherent (9.0/2.0)
| | | | | | InitialWHOstage = 1
| | | | | | | MaritalStatus = Never-Married: NonAdherent (12.0/4.0)
| | | | | | | MaritalStatus = Divorced: Adherent (2.0)
| | | | | | | MaritalStatus = Married: Adherent (16.0/4.0)
| | | | | | | MaritalStatus = Widowed: Adherent (2.0)
| | | | | | | MaritalStatus = Living-Together: Adherent (0.0)
| | | | | | InitialWHOstage = 2
| | | | | | | MaritalStatus = Never-Married: Adherent (31.0/14.0)
| | | | | | | MaritalStatus = Divorced: NonAdherent (4.0)
| | | | | | | MaritalStatus = Married: Adherent (42.0/14.0)
| | | | | | | MaritalStatus = Widowed: NonAdherent (8.0/4.0)
| | | | | | | MaritalStatus = Living-Together: NonAdherent (2.0)
| | | | | | InitialWHOstage = 3
| | | | | | | AgeatEnrolment <= 37
| | | | | | | | HasPastOI = Yes: Adherent (10.0/4.0)
| | | | | | | | HasPastOI = No: NonAdherent (46.0/12.0)
| | | | | | | AgeatEnrolment > 37: Adherent (18.0/4.0)
| | | | | | InitialWHOstage = 4
| | | | | | | MaritalStatus = Never-Married: NonAdherent (16.0/8.0)
| | | | | | | MaritalStatus = Divorced: NonAdherent (4.0/2.0)
| | | | | | | MaritalStatus = Married: NonAdherent (11.0/2.0)
| | | | | | | MaritalStatus = Widowed: Adherent (5.0/1.0)
| | | | | | | MaritalStatus = Living-Together: Adherent (1.0)
| | | | | HasEmergencyContact = No: Adherent (42.0/12.0)
| | | | HasDependent = No
| | | | | AgeatEnrolment <= 21: NonAdherent (11.0/2.0)
| | | | | AgeatEnrolment > 21: Adherent (189.0)
| | | InitialRegimenCode = 1e
| | | | HasDependent = Yes
| | | | | AgeatEnrolment <= 39
| | | | | | PhoneAddressForTracingExist = Yes
| | | | | | | MonthsonART <= 105: Adherent (379.0/113.0)
| | | | | | | MonthsonART > 105
| | | | | | | | FunctionalStatus = Working
| | | | | | | | | MaritalStatus = Never-Married: Adherent (50.0/23.0)
| | | | | | | | | MaritalStatus = Divorced: NonAdherent (11.0/5.0)
| | | | | | | | | MaritalStatus = Married
| | | | | | | | | | Sex = M: Adherent (39.0/16.0)
| | | | | | | | | | Sex = F: NonAdherent (57.0/25.0)
| | | | | | | | | MaritalStatus = Widowed: Adherent (17.0/4.0)
| | | | | | | | | MaritalStatus = Living-Together: NonAdherent (6.0/2.0)
| | | | | | | | FunctionalStatus = Bedridden
| | | | | | | | | MonthsonART <= 119: NonAdherent (12.0/2.0)
| | | | | | | | | MonthsonART > 119: Adherent (19.0/8.0)
```

```
| | | | | | | | FunctionalStatus = Ambulatory: NonAdherent (14.0/4.0)
| | | | | | PhoneAddressForTracingExist = No: Adherent (32.0/6.0)
| | | | | AgeatEnrolment > 39: Adherent (53.0/7.0)
| | | | HasDependent = No
| | | | | AgeatEnrolment <= 21: NonAdherent (17.0/2.0)
| | | | | AgeatEnrolment > 21: Adherent (184.0/1.0)
| | | InitialRegimenCode = 1f
| | | | HasEmergencyContact = Yes
| | | | | FunctionalStatus = Working: NonAdherent (94.0/42.0)
| | | | | FunctionalStatus = Bedridden: Adherent (17.0/8.0)
| | | | | FunctionalStatus = Ambulatory: Adherent (2.0)
| | | | HasEmergencyContact = No: Adherent (50.0/5.0)
| | | InitialRegimenCode = 1g: Adherent (14.0/4.0)
| | MonthsonART > 201
| | | PhoneAddressForTracingExist = Yes
| | | | InitialRegimenCode = 1a: NonAdherent (1.0)
| | | | InitialRegimenCode = 1a30
| | | | | HasEmergencyContact = Yes
| | | | | | AgeatEnrolment <= 29: NonAdherent (16.0/4.0)
| | | | | | AgeatEnrolment > 29
| | | | | | | Sex = M
| | | | | | | | AgeatEnrolment <= 36: NonAdherent (17.0/7.0)
| | | | | | | | AgeatEnrolment > 36: Adherent (10.0/2.0)
| | | | | | | Sex = F: Adherent (26.0/6.0)
| | | | | HasEmergencyContact = No
| | | | | | MonthsonART <= 216: Adherent (20.0/9.0)
| | | | | | MonthsonART > 216: NonAdherent (16.0/2.0)
| | | | InitialRegimenCode = 1a40
| | | | | HasDependent = Yes
| | | | | | HasPastOI = Yes: NonAdherent (14.0/6.0)
| | | | | | HasPastOI = No: Adherent (23.0/7.0)
| | | | | HasDependent = No: NonAdherent (11.0/1.0)
| | | | InitialRegimenCode = 1b: Adherent (1.0)
| | | | InitialRegimenCode = 1b30
| | | | | Religion = NULL: NonAdherent (15.0/4.0)
| | | | | Religion = Protestant: NonAdherent (1.0)
| | | | | Religion = Orthodox: Adherent (15.0/4.0)
| | | | InitialRegimenCode = 1b40: Adherent (7.0/3.0)
| | | | InitialRegimenCode = 1c
| | | | | HasEmergencyContact = Yes
| | | | | | ARTInitationDuration = Beyond 2 Weeks: NonAdherent (49.0/16.0)
| | | | | | ARTInitationDuration = Same Day
| | | | | | | AgeatEnrolment <= 33: NonAdherent (11.0/4.0)
| | | | | | | AgeatEnrolment > 33: Adherent (11.0/2.0)
| | | | | | ARTInitationDuration = Within 1 Week: Adherent (11.0/4.0)
| | | | | HasEmergencyContact = No: Adherent (30.0/3.0)
| | | | InitialRegimenCode = 1d
| | | | | Religion = NULL: Adherent (9.0)
| | | | | Religion = Protestant: NonAdherent (12.0/4.0)
| | | | | Religion = Orthodox
| | | | | | MaritalStatus = Never-Married: NonAdherent (21.0/9.0)
| | | | | | MaritalStatus = Divorced: Adherent (3.0)
| | | | | | MaritalStatus = Married: NonAdherent (15.0/4.0)
| | | | | | MaritalStatus = Widowed: Adherent (4.0/1.0)
| | | | | | MaritalStatus = Living-Together: NonAdherent (2.0/1.0)
| | | | | Religion = Muslim: Adherent (6.0/2.0)
| | | | | Religion = Other: Adherent (0.0)
| | | | | Religion = Catholic: NonAdherent (1.0)
| | | | | Religion = PROTESTANT: Adherent (0.0)
| | | | | Religion = ORTHODOX: Adherent (0.0)
| | | | | Religion = MUSLIM: Adherent (0.0)
| | | | | Religion = CATHOLIC: Adherent (0.0)
```

```
|  |  |  |  InitialRegimenCode = 1e: Adherent (3.0)
|  |  |  |  InitialRegimenCode = 1f: Adherent (2.0)
|  |  |  |  InitialRegimenCode = 1g: Adherent (2.0)
|  |  |  |  InitialRegimenCode = 1h: Adherent (0.0)
|  |  |  PhoneAddressForTracingExist = No: Adherent (1062.0/136.0)
|  AgeatEnrolment > 40
|  |  AgeatEnrolment <= 50
|  |  |  HasEmergencyContact = Yes
|  |  |  |  MonthsonART <= 122: Adherent (255.0/63.0)
|  |  |  |  MonthsonART > 122
|  |  |  |  |  InitialWeight <= 43: NonAdherent (16.0/2.0)
|  |  |  |  |  InitialWeight > 43
|  |  |  |  |  |  Religion = NULL
|  |  |  |  |  |  |  InitialRegimenCode = 1a: Adherent (0.0)
|  |  |  |  |  |  |  InitialRegimenCode = 1a30: Adherent (4.0/1.0)
|  |  |  |  |  |  |  InitialRegimenCode = 1a40: NonAdherent (4.0)
|  |  |  |  |  |  Religion = Protestant: Adherent (14.0/3.0)
|  |  |  |  |  |  Religion = Orthodox
|  |  |  |  |  |  |  MaritalStatus = Never-Married
|  |  |  |  |  |  |  |  MonthsonART <= 191: Adherent (20.0/6.0)
|  |  |  |  |  |  |  |  MonthsonART > 191: NonAdherent (11.0/3.0)
|  |  |  |  |  |  |  MaritalStatus = Divorced: NonAdherent (20.0/8.0)
|  |  |  |  |  |  |  MaritalStatus = Married: NonAdherent (120.0/56.0)
|  |  |  |  |  |  |  MaritalStatus = Widowed: Adherent (41.0/15.0)
|  |  |  |  |  |  |  MaritalStatus = Living-Together: Adherent (6.0/2.0)
|  |  |  |  |  |  Religion = Muslim: Adherent (10.0/4.0)
|  |  |  |  |  |  Religion = Other: NonAdherent (2.0)
|  |  |  |  |  |  Religion = Catholic: Adherent (0.0)
|  |  |  |  |  |  Religion = PROTESTANT: Adherent (0.0)
|  |  |  |  |  |  Religion = ORTHODOX: Adherent (0.0)
|  |  |  |  |  |  Religion = MUSLIM: NonAdherent (1.0)
|  |  |  |  |  |  Religion = CATHOLIC: Adherent (0.0)
|  |  |  HasEmergencyContact = No
|  |  |  |  MonthsonART <= 153
|  |  |  |  |  InitialRegimenCode = 1a: NonAdherent (0.0)
|  |  |  |  |  InitialRegimenCode = 1a30: NonAdherent (0.0)
|  |  |  |  |  InitialRegimenCode = 1a40: NonAdherent (0.0)
|  |  |  |  |  InitialRegimenCode = 1b: NonAdherent (0.0)
|  |  |  |  |  InitialRegimenCode = 1b30: NonAdherent (2.0)
|  |  |  |  |  InitialRegimenCode = 1b40: NonAdherent (0.0)
|  |  |  |  |  InitialRegimenCode = 1c: NonAdherent (10.0)
|  |  |  |  |  InitialRegimenCode = 1d: NonAdherent (33.0/8.0)
|  |  |  |  |  InitialRegimenCode = 1e
|  |  |  |  |  |  Religion = NULL
|  |  |  |  |  |  |  PhoneAddressForTracingExist = Yes
|  |  |  |  |  |  |  |  InitialWeight <= 80: NonAdherent (102.0/29.0)
|  |  |  |  |  |  |  |  InitialWeight > 80: Adherent (11.0/2.0)
|  |  |  |  |  |  |  PhoneAddressForTracingExist = No
|  |  |  |  |  |  |  |  InitialWHOstage = 0: NonAdherent (6.0/2.0)
|  |  |  |  |  |  |  |  InitialWHOstage = 1: NonAdherent (15.0/6.0)
|  |  |  |  |  |  |  |  InitialWHOstage = 2: Adherent (25.0/8.0)
|  |  |  |  |  |  |  |  InitialWHOstage = 3: Adherent (11.0/1.0)
|  |  |  |  |  |  |  |  InitialWHOstage = 4: NonAdherent (2.0/1.0)
|  |  |  |  |  |  |  |  InitialWHOstage = 5: Adherent (0.0)
|  |  |  |  |  |  |  |  InitialWHOstage = 6: Adherent (0.0)
|  |  |  |  |  |  |  |  InitialWHOstage = 7: Adherent (0.0)
|  |  |  |  |  |  Religion = Protestant: Adherent (2.0)
|  |  |  |  |  |  Religion = Orthodox: Adherent (10.0/4.0)
|  |  |  |  |  |  Religion = Muslim: Adherent (3.0)
|  |  |  |  |  |  Religion = Other: NonAdherent (0.0)
|  |  |  |  |  |  Religion = Catholic: NonAdherent (0.0)
|  |  |  |  |  |  Religion = PROTESTANT: NonAdherent (0.0)
```

```
| | | | | | Religion = ORTHODOX: NonAdherent (0.0)
| | | | | | Religion = MUSLIM: NonAdherent (0.0)
| | | | | | Religion = CATHOLIC: NonAdherent (0.0)
| | | | | InitialRegimenCode = 1f
| | | | | | PhoneAddressForTracingExist = Yes: NonAdherent (16.0/4.0)
| | | | | | PhoneAddressForTracingExist = No: Adherent (11.0/4.0)
| | | | MonthsonART > 153: NonAdherent (595.0/71.0)
| | AgeatEnrolment > 50: Adherent (387.0/96.0)
EducationalLevelAttained = No education
| MonthsonART <= 27: Adherent (40.0/6.0)
| MonthsonART > 27
| | InitialRegimenLine = 1
| | | ARTInitationDuration = Beyond 2 Weeks
| | | | InitialRegimenCode = 1a: Adherent (0.0)
| | | | InitialRegimenCode = 1a30
| | | | | InitialWeight <= 48.5: NonAdherent (17.0/5.0)
| | | | | InitialWeight > 48.5: Adherent (29.0/8.0)
| | | | InitialRegimenCode = 1a40: NonAdherent (2.0/1.0)
| | | | InitialRegimenCode = 1b: Adherent (0.0)
| | | | InitialRegimenCode = 1b30
| | | | | MonthsonART <= 179: NonAdherent (16.0/4.0)
| | | | | MonthsonART > 179: Adherent (13.0/4.0)
| | | | InitialRegimenCode = 1b40: NonAdherent (2.0/1.0)
| | | | InitialRegimenCode = 1c
| | | | | MaritalStatus = Never-Married: Adherent (7.0/1.0)
| | | | | MaritalStatus = Divorced: NonAdherent (11.0/5.0)
| | | | | MaritalStatus = Married: NonAdherent (18.0/8.0)
| | | | | MaritalStatus = Widowed: NonAdherent (13.0/3.0)
| | | | | MaritalStatus = Living-Together: Adherent (3.0/1.0)
| | | | InitialRegimenCode = 1d: Adherent (31.0/12.0)
| | | | InitialRegimenCode = 1e
| | | | | Sex = M: NonAdherent (22.0/8.0)
| | | | | Sex = F: Adherent (93.0/32.0)
| | | | InitialRegimenCode = 1f: NonAdherent (14.0/5.0)
| | | | InitialRegimenCode = 1g: NonAdherent (2.0/1.0)
| | | ARTInitationDuration = Same Day: NonAdherent (53.0/20.0)
| | | ARTInitationDuration = Within 1 Week
| | | | AgeatEnrolment <= 40: NonAdherent (27.0/7.0)
| | | | AgeatEnrolment > 40: Adherent (28.0/12.0)
| | InitialRegimenLine = 2: Adherent (29.0/9.0)
| | InitialRegimenLine = 3: Adherent (1.0)
EducationalLevelAttained = Primary
| AgeatEnrolment <= 40
| | HasDependent = Yes
| | | HasEmergencyContact = Yes
| | | | Occupation = NULL
| | | | | Religion = NULL: NonAdherent (169.0/58.0)
| | | | | Religion = Protestant
| | | | | | AgeatEnrolment <= 29: NonAdherent (25.0/3.0)
| | | | | | AgeatEnrolment > 29: Adherent (52.0/23.0)
| | | | | Religion = Orthodox
| | | | | | Sex = M
| | | | | | | InitialRegimenCode = 1a: NonAdherent (0.0)
| | | | | | | InitialRegimenCode = 1a30: NonAdherent (30.0/9.0)
| | | | | | | InitialRegimenCode = 1a40: NonAdherent (5.0/2.0)
| | | | | | | InitialRegimenCode = 1b: NonAdherent (0.0)
| | | | | | | InitialRegimenCode = 1b30
| | | | | | | | AgeatEnrolment <= 29: NonAdherent (11.0/3.0)
| | | | | | | | AgeatEnrolment > 29: Adherent (11.0/4.0)
| | | | | | | InitialRegimenCode = 1b40: Adherent (3.0/1.0)
| | | | | | | InitialRegimenCode = 1c
| | | | | | | | MaritalStatus = Never-Married: NonAdherent (17.0/7.0)
```

```
| | | | | | | | MaritalStatus = Divorced: NonAdherent (2.0)
| | | | | | | | MaritalStatus = Married: Adherent (20.0/7.0)
| | | | | | | | MaritalStatus = Widowed: NonAdherent (2.0/1.0)
| | | | | | | | MaritalStatus = Living-Together: NonAdherent (1.0)
| | | | | | | InitialRegimenCode = 1d: NonAdherent (49.0/17.0)
| | | | | | | InitialRegimenCode = 1e
| | | | | | | | MaritalStatus = Never-Married: NonAdherent (34.0/8.0)
| | | | | | | | MaritalStatus = Divorced: NonAdherent (3.0/1.0)
| | | | | | | | MaritalStatus = Married
| | | | | | | | | AgeatEnrolment <= 35: Adherent (12.0/1.0)
| | | | | | | | | AgeatEnrolment > 35
| | | | | | | | | | InitialWeight <= 57: Adherent (12.0/2.0)
| | | | | | | | | | InitialWeight > 57: NonAdherent (12.0/4.0)
| | | | | | | | MaritalStatus = Widowed: Adherent (2.0)
| | | | | | | | MaritalStatus = Living-Together: NonAdherent (0.0)
| | | | | | | InitialRegimenCode = 1f: NonAdherent (8.0/1.0)
| | | | | | | InitialRegimenCode = 1g: NonAdherent (0.0)
| | | | | | Sex = F
| | | | | | | PhoneAddressForTracingExist = Yes
| | | | | | | | InitialRegimenCode = 1a: Adherent (0.0)
| | | | | | | | InitialRegimenCode = 1a30
| | | | | | | | | MaritalStatus = Never-Married: NonAdherent (12.0/3.0)
| | | | | | | | | MaritalStatus = Divorced: Adherent (7.0/1.0)
| | | | | | | | | MaritalStatus = Married
| | | | | | | | | | AgeatEnrolment <= 31: Adherent (14.0/6.0)
| | | | | | | | | | AgeatEnrolment > 31: NonAdherent (10.0/1.0)
| | | | | | | | | MaritalStatus = Widowed: NonAdherent (8.0/3.0)
| | | | | | | | | MaritalStatus = Living-Together: Adherent (3.0)
| | | | | | | | InitialRegimenCode = 1a40: NonAdherent (6.0/1.0)
| | | | | | | | InitialRegimenCode = 1b: Adherent (0.0)
| | | | | | | | InitialRegimenCode = 1b30
| | | | | | | | | InitialWHOstage = 0: NonAdherent (0.0)
| | | | | | | | | InitialWHOstage = 1: Adherent (1.0)
| | | | | | | | | InitialWHOstage = 2: NonAdherent (8.0/3.0)
| | | | | | | | | InitialWHOstage = 3
| | | | | | | | | | InitialWeight <= 44: NonAdherent (11.0/4.0)
| | | | | | | | | | InitialWeight > 44: Adherent (13.0/4.0)
| | | | | | | | | InitialWHOstage = 4: NonAdherent (12.0/3.0)
| | | | | | | | | InitialWHOstage = 5: NonAdherent (0.0)
| | | | | | | | | InitialWHOstage = 6: NonAdherent (0.0)
| | | | | | | | | InitialWHOstage = 7: NonAdherent (0.0)
| | | | | | | | InitialRegimenCode = 1b40: NonAdherent (3.0/1.0)
| | | | | | | | InitialRegimenCode = 1c
| | | | | | | | | HasPastOI = Yes: Adherent (13.0/3.0)
| | | | | | | | | HasPastOI = No: NonAdherent (97.0/44.0)
| | | | | | | | InitialRegimenCode = 1d
| | | | | | | | | ARTInitationDuration = Beyond 2 Weeks: NonAdherent (47.0/22.0)
| | | | | | | | | ARTInitationDuration = Same Day: Adherent (2.0)
| | | | | | | | | ARTInitationDuration = Within 1 Week: Adherent (12.0/5.0)
| | | | | | | | InitialRegimenCode = 1e
| | | | | | | | | FunctionalStatus = Working: Adherent (148.0/56.0)
| | | | | | | | | FunctionalStatus = Bedridden: Adherent (10.0/3.0)
| | | | | | | | | FunctionalStatus = Ambulatory: NonAdherent (8.0/1.0)
| | | | | | | | InitialRegimenCode = 1f
| | | | | | | | | InitialWHOstage = 0: NonAdherent (6.0/2.0)
| | | | | | | | | InitialWHOstage = 1: Adherent (11.0/5.0)
| | | | | | | | | InitialWHOstage = 2: Adherent (18.0/7.0)
| | | | | | | | | InitialWHOstage = 3: Adherent (9.0/4.0)
| | | | | | | | | InitialWHOstage = 4: NonAdherent (3.0)
| | | | | | | | InitialRegimenCode = 2f: Adherent (4.0/1.0)
| | | | | | | | InitialRegimenCode = 2g: NonAdherent (2.0)
| | | | | | | | InitialRegimenCode = 2h: Adherent (6.0/1.0)
```

```
| | | | | | | | InitialRegimenCode = Oth2A: Adherent (9.0/4.0)
| | | | | | | | InitialRegimenCode = Oth3A: Adherent (0.0)
| | | | | | | | InitialRegimenCode = 1k: Adherent (0.0)
| | | | | | | | InitialRegimenCode = 3a: Adherent (0.0)
| | | | | | | | InitialRegimenCode = 4i: Adherent (0.0)
| | | | | | | | InitialRegimenCode = Oth1C: Adherent (0.0)
| | | | | | | PhoneAddressForTracingExist = No: Adherent (27.0/6.0)
| | | | | Religion = Muslim
| | | | | | MaritalStatus = Never-Married: NonAdherent (35.0/11.0)
| | | | | | MaritalStatus = Divorced: Adherent (8.0/2.0)
| | | | | | MaritalStatus = Married
| | | | | | | AgeatEnrolment <= 31: Adherent (31.0/12.0)
| | | | | | | AgeatEnrolment > 31: NonAdherent (27.0/9.0)
| | | | | | MaritalStatus = Widowed: NonAdherent (2.0/1.0)
| | | | | | MaritalStatus = Living-Together: Adherent (3.0/1.0)
| | | | | Religion = Other: NonAdherent (5.0/1.0)
| | | | | Religion = Catholic: Adherent (6.0/2.0)
| | | | | Religion = PROTESTANT: NonAdherent (0.0)
| | | | | Religion = ORTHODOX: NonAdherent (0.0)
| | | | | Religion = MUSLIM: NonAdherent (0.0)
| | | | | Religion = CATHOLIC: NonAdherent (0.0)
| | | | Occupation = UNEMPLOYED: NonAdherent (19.0/4.0)
| | | | Occupation = HOUSEWIFE: Adherent (31.0/9.0)
| | | | Occupation = Draughtspersons: NonAdherent (0.0)
| | | | Occupation = OTHER
| | | | | Sex = M: Adherent (13.0/4.0)
| | | | | Sex = F: NonAdherent (18.0/7.0)
| | | | Occupation = TEACHER: NonAdherent (0.0)
| | | HasEmergencyContact = No
| | | | Religion = NULL: NonAdherent (145.0/14.0)
| | | | Religion = Protestant: NonAdherent (12.0/5.0)
| | | | Religion = Orthodox
| | | | | ARTInitationDuration = Beyond 2 Weeks
| | | | | | MaritalStatus = Never-Married: NonAdherent (16.0/8.0)
| | | | | | MaritalStatus = Divorced: NonAdherent (17.0/5.0)
| | | | | | MaritalStatus = Married: Adherent (26.0/10.0)
| | | | | | MaritalStatus = Widowed: Adherent (17.0/7.0)
| | | | | | MaritalStatus = Living-Together: NonAdherent (3.0)
| | | | | ARTInitationDuration = Same Day: NonAdherent (18.0/5.0)
| | | | | ARTInitationDuration = Within 1 Week: NonAdherent (12.0/2.0)
| | | | Religion = Muslim: Adherent (24.0/9.0)
| | | | Religion = Other: NonAdherent (0.0)
| | | | Religion = Catholic: NonAdherent (1.0)
| | | | Religion = PROTESTANT: NonAdherent (0.0)
| | | | Religion = ORTHODOX: NonAdherent (0.0)
| | | | Religion = MUSLIM: NonAdherent (0.0)
| | | | Religion = CATHOLIC: NonAdherent (0.0)
| | HasDependent = No: NonAdherent (1667.0/15.0)
| AgeatEnrolment > 40
| | AgeatEnrolment <= 50
| | | HasDependent = Yes
| | | | InitialRegimenCode = 1a: Adherent (1.0)
| | | | InitialRegimenCode = 1a30
| | | | | Sex = M: Adherent (29.0/8.0)
| | | | | Sex = F: NonAdherent (13.0/4.0)
| | | | InitialRegimenCode = 1a40: NonAdherent (12.0/4.0)
| | | | InitialRegimenCode = 1b: Adherent (0.0)
| | | | InitialRegimenCode = 1b30
| | | | | InitialWeight <= 47: NonAdherent (10.0/2.0)
| | | | | InitialWeight > 47: Adherent (26.0/8.0)
| | | | InitialRegimenCode = 1b40: Adherent (2.0)
| | | | InitialRegimenCode = 1c: Adherent (77.0/24.0)
```

```
|   |   |   |   InitialRegimenCode = 1d
|   |   |   |   |   ARTInitationDuration = Beyond 2 Weeks: Adherent (53.0/22.0)
|   |   |   |   |   ARTInitationDuration = Same Day: NonAdherent (3.0/1.0)
|   |   |   |   |   ARTInitationDuration = Within 1 Week: NonAdherent (14.0/5.0)
|   |   |   |   InitialRegimenCode = 1e
|   |   |   |   |   Religion = NULL: NonAdherent (10.0/4.0)
|   |   |   |   |   Religion = Protestant: NonAdherent (4.0/2.0)
|   |   |   |   |   Religion = Orthodox
|   |   |   |   |   |   InitialWeight <= 59.5
|   |   |   |   |   |   |   MonthsonART <= 62: Adherent (13.0/2.0)
|   |   |   |   |   |   |   MonthsonART > 62: NonAdherent (35.0/11.0)
|   |   |   |   |   |   InitialWeight > 59.5: Adherent (51.0/9.0)
|   |   |   |   |   Religion = Muslim: Adherent (14.0/4.0)
|   |   |   |   |   Religion = Other: NonAdherent (1.0)
|   |   |   |   |   Religion = Catholic: NonAdherent (1.0)
|   |   |   |   |   Religion = PROTESTANT: Adherent (0.0)
|   |   |   |   |   Religion = ORTHODOX: Adherent (0.0)
|   |   |   |   |   Religion = MUSLIM: Adherent (0.0)
|   |   |   |   |   Religion = CATHOLIC: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 1f: Adherent (6.0/2.0)
|   |   |   |   InitialRegimenCode = 1g: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 1h: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 1j: Adherent (31.0/4.0)
|   |   |   |   InitialRegimenCode = 2a: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 2e: NonAdherent (1.0)
|   |   |   |   InitialRegimenCode = 2f: NonAdherent (3.0/1.0)
|   |   |   |   InitialRegimenCode = 2g: Adherent (3.0/1.0)
|   |   |   |   InitialRegimenCode = 2h: Adherent (5.0/2.0)
|   |   |   |   InitialRegimenCode = 2i: Adherent (1.0)
|   |   |   |   InitialRegimenCode = 4a: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 4b: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 4c: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 4d: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 4e: Adherent (0.0)
|   |   |   |   InitialRegimenCode = Oth1A: Adherent (1.0)
|   |   |   |   InitialRegimenCode = Oth2A: NonAdherent (5.0/1.0)
|   |   |   |   InitialRegimenCode = Oth3A: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 1k: Adherent (1.0)
|   |   |   |   InitialRegimenCode = 3a: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 4i: Adherent (0.0)
|   |   |   |   InitialRegimenCode = Oth1C: Adherent (0.0)
|   |   |   HasDependent = No: Adherent (434.0/1.0)
|   |   AgeatEnrolment > 50
|   |   |   HasDependent = Yes
|   |   |   |   InitialRegimenCode = 1a: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 1a30: NonAdherent (17.0/7.0)
|   |   |   |   InitialRegimenCode = 1a40: NonAdherent (4.0/1.0)
|   |   |   |   InitialRegimenCode = 1b: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 1b30: NonAdherent (10.0/5.0)
|   |   |   |   InitialRegimenCode = 1b40: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 1c: NonAdherent (32.0/12.0)
|   |   |   |   InitialRegimenCode = 1d: Adherent (31.0/14.0)
|   |   |   |   InitialRegimenCode = 1e
|   |   |   |   |   InitialWeight <= 73: Adherent (61.0/20.0)
|   |   |   |   |   InitialWeight > 73: NonAdherent (11.0/2.0)
|   |   |   |   InitialRegimenCode = 1f: Adherent (7.0/3.0)
|   |   |   |   InitialRegimenCode = 1g: Adherent (9.0/4.0)
|   |   |   |   InitialRegimenCode = 1h: NonAdherent (2.0/1.0)
|   |   |   |   InitialRegimenCode = 1j: Adherent (11.0)
|   |   |   |   InitialRegimenCode = 2a: Adherent (0.0)
|   |   |   |   InitialRegimenCode = 2e: NonAdherent (1.0)
|   |   |   |   InitialRegimenCode = 2f: Adherent (1.0)
```

```
| | | | InitialRegimenCode = 2g: Adherent (0.0)
| | | | InitialRegimenCode = 2h: NonAdherent (2.0)
| | | | InitialRegimenCode = 2i: NonAdherent (1.0)
| | | | InitialRegimenCode = 4a: Adherent (0.0)
| | | | InitialRegimenCode = 4b: Adherent (0.0)
| | | | InitialRegimenCode = 4c: Adherent (0.0)
| | | | InitialRegimenCode = 4d: Adherent (0.0)
| | | | InitialRegimenCode = 4e: Adherent (0.0)
| | | | InitialRegimenCode = Oth1A: Adherent (1.0)
| | | | InitialRegimenCode = Oth2A: Adherent (2.0)
| | | | InitialRegimenCode = Oth3A: Adherent (0.0)
| | | | InitialRegimenCode = 1k: Adherent (0.0)
| | | | InitialRegimenCode = 3a: Adherent (0.0)
| | | | InitialRegimenCode = 4i: Adherent (0.0)
| | | | InitialRegimenCode = Oth1C: Adherent (0.0)
| | | HasDependent = No
| | | | AgeatEnrolment <= 69: NonAdherent (266.0/27.0)
| | | | AgeatEnrolment > 69
| | | | | MonthsonART <= 191: NonAdherent (20.0/5.0)
| | | | | MonthsonART > 191: Adherent (11.0/2.0)
EducationalLevelAttained = Tertiary
| Region = Addis Ababa
| | HasDependent = Yes
| | | MonthsonART <= 102: Adherent (410.0/130.0)
| | | MonthsonART > 102
| | | | ARTInitationDuration = Beyond 2 Weeks
| | | | | InitialRegimenCode = 1a: Adherent (0.0)
| | | | | InitialRegimenCode = 1c
| | | | | | InitialWHOstage = 0: NonAdherent (5.0/2.0)
| | | | | | InitialWHOstage = 1: NonAdherent (18.0/8.0)
| | | | | | InitialWHOstage = 2: Adherent (47.0/20.0)
| | | | | | InitialWHOstage = 3
| | | | | | | MonthsonART <= 187: NonAdherent (15.0/5.0)
| | | | | | | MonthsonART > 187: Adherent (11.0/3.0)
| | | | | | InitialWHOstage = 4: NonAdherent (12.0/4.0)
| | | | | InitialRegimenCode = 1d: Adherent (118.0/44.0)
| | | | | InitialRegimenCode = 1e
| | | | | | AgeatEnrolment <= 37
| | | | | | | MaritalStatus = Never-Married: NonAdherent (22.0/5.0)
| | | | | | | MaritalStatus = Divorced: NonAdherent (4.0/2.0)
| | | | | | | MaritalStatus = Married
| | | | | | | | InitialWeight <= 55.5: NonAdherent (13.0/5.0)
| | | | | | | | InitialWeight > 55.5: Adherent (19.0/7.0)
| | | | | | | MaritalStatus = Widowed: NonAdherent (4.0/2.0)
| | | | | | | MaritalStatus = Living-Together: NonAdherent (0.0)
| | | | | | AgeatEnrolment > 37: Adherent (56.0/16.0)
| | | | | InitialRegimenCode = 1f: Adherent (29.0/10.0)
| | | | | InitialRegimenCode = Oth2A: Adherent (3.0/1.0)
| | | | ARTInitationDuration = Same Day: NonAdherent (121.0/50.0)
| | | | ARTInitationDuration = Within 1 Week: Adherent (117.0/55.0)
| | HasDependent = No
| | | AgeatEnrolment <= 28: NonAdherent (57.0/6.0)
| | | AgeatEnrolment > 28
| | | | Religion = NULL: Adherent (5.0)
| | | | Religion = Protestant: Adherent (0.0)
| | | | Religion = Orthodox: Adherent (0.0)
| | | | Religion = Muslim: Adherent (0.0)
| | | | Religion = Other: Adherent (1.0)
| | | | Religion = Catholic: Adherent (0.0)
| | | | Religion = PROTESTANT: Adherent (17.0/6.0)
| | | | Religion = ORTHODOX: NonAdherent (47.0/20.0)
| | | | Religion = MUSLIM: Adherent (3.0/1.0)
```

```
| | | | Religion = CATHOLIC: Adherent (3.0/1.0)
| Region = Oromia
| | MonthsonART <= 97: Adherent (19.0/4.0)
| | MonthsonART > 97: NonAdherent (20.0/6.0)
| Region = Amhara: Adherent (8.0)
| Region = Afar: Adherent (0.0)
| Region = SNNPR: NonAdherent (2.0)
| Region = Harari: NonAdherent (3.0)
| Region = Tigray: NonAdherent (3.0/1.0)
| Region = Benshuagul: Adherent (0.0)
| Region = Somali: NonAdherent (1.0)
EducationalLevelAttained = SECONDARY
| HasPastOI = Yes: NonAdherent (25.0/6.0)
| HasPastOI = No
| | Sex = M
| | | ARTInitationDuration = Beyond 2 Weeks: NonAdherent (78.0/26.0)
| | | ARTInitationDuration = Same Day: Adherent (7.0/1.0)
| | | ARTInitationDuration = Within 1 Week: Adherent (20.0/9.0)
| | Sex = F: Adherent (90.0/31.0)
EducationalLevelAttained = PRIMARY: NonAdherent (169.0/66.0)
EducationalLevelAttained = TERTIARY: NonAdherent (5.0/2.0)
EducationalLevelAttained = NO EDUCATION
| MonthsonART <= 75: Adherent (16.0/5.0)
| MonthsonART > 75: NonAdherent (21.0/4.0)

Number of Leaves  :         714

Size of the tree :      871

Time taken to build model: 0.16 seconds

=== Evaluation on test split ===

Time taken to test model on training split: 0 seconds

=== Summary ===

Correctly Classified Instances        1816               71.4398 %
Incorrectly Classified Instances       726               28.5602 %
Kappa statistic                          0.4291
Mean absolute error                      0.3386
Root mean squared error                  0.4309
Relative absolute error                 67.7156 %
Root relative squared error             86.1515 %
Total Number of Instances             2542

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               0.689    0.259    0.731      0.689    0.709      0.430    0.795     0.805     NonAdherent
               0.741    0.311    0.699      0.741    0.719      0.430    0.795     0.760     Adherent
Weighted Avg.  0.714    0.285    0.715      0.714    0.714      0.430    0.795     0.783

=== Confusion Matrix ===

   a   b   <-- classified as
 885 400 |   a = NonAdherent
 326 931 |   b = Adherent
```