`

PhD Thesis

Title: Utilisation of maternal, newborn and child healthcare services in three sub-Saharan African countries (DRC, Kenya, and Tanzania) using Demographic Health Surveys data from 2007-2016: Application of Generalised Structural Equation and Machine Learning Models.

By Chenai Mlandu
Student number 1155482

Supervisors: Professor Eustasius Musenge
Dr Zvifadzo Matsena-Zingoni

A Thesis submitted to the Faculty of Health Sciences, University of Witwatersrand, Johannesburg, in fulfilment for requirements for a degree of Doctor of Philosophy in Public Health (by publication).
Johannesburg, 2024

CANDIDATE DECLARATION

I Chenai Mlandu declare that this thesis is entirely my own work. It is being submitted for a Doctor of Philosophy degree in Public Health at the University of Witwatersrand in Johannesburg. It has not been submitted for any degree or examination at any other University or institution.

Signature of Candidate

CMlandu on the 24 of October 2024

DEDICATION

This thesis is dedicated to my mother Joyce Mlandu, my brother Audwell Mlandu and the rest of my family. I am deeply grateful for your love and support.

PREFACE

My interest in epidemiology stems from my time at the National University of Science and Technology, where I earned a Bachelor of Science Honors degree in Environmental Science and Health. The curriculum was broad, covering environmental health, pollution control, environmental management, toxicology, environmental engineering, epidemiology, sanitation and safety, and food laws. Because of my interest in epidemiology, I pursued additional studies in MSC Infectious Disease Epidemiology. During this program, I learned more about applying epidemiology and biostatistics tools to investigate disease occurrence and patterns in populations. I also developed a strong interest in research.

After completing my MSC, I began working as a research assistant in the School of Public Health on maternal and child health projects. During this work, I became interested in pursuing a PhD in maternal and child health. I was also interested in analysing data using a variety of quantitative statistical techniques, including GSEM and Machine Learning. Aside from expanding my statistical knowledge and skills, pursuing a PhD through publication allowed me to learn research and publication techniques. As a result, I conducted doctoral research assessing the utilisation of maternal, newborn and child healthcare services and associated neonatal outcomes in three sub-Saharan African countries using Demographic Health Surveys data from 2007-2016 using Generalised Structural Equation and Machine Learning Models.

Conferences and Scholarships

Conferences

Attended and presented research findings at the 16th Early Scientist Career Convention organised by the South African Medical Research Council (SAMRC) in Capetown in 2022.

1. Oral Presentation. Mlandu C, Matsena-Zingoni Z, Musenge E (2023). Predicting the drop out from the maternal, newborn and child healthcare continuum in three East African Community countries: Application of machine learning models. Presented at the 16th Early Scientist Career Convention, SAMRC, Capetown, South Africa, 2022.

Scholarships

Held scholarships during the course of the PhD

1. University of Witwatersrand Postgraduate (PhD) Merit Award (2019-2021)

2. Bongani Mayosi National Health Scholars (SAMRC) Award (2021-2023)

My contribution to the THESIS

The research was written up in a series of paper publications during this PhD. The papers are presented in full in chapter 4 to 7.

Publications

In paper 1, the study sought to assess the trends and determinants of late ANC initiation in the DRC, Kenya and Tanzania from 2007 to 2016.

1. Mlandu C, Matsena-Zingoni Z, Musenge E (2022) Trends and determinants of late antenatal care initiation in three East African countries, 2007–2016: A population based cross-sectional analysis. PLOS Glob Public Health 2(8): e0000534.

PhD student: Responsible for the manuscript's conceptualisation, which included defining the objectives and methodology to be used, performing data management (including data cleaning, coding, and manipulation), data analysis (including mapping and multilevel logistic regression), drafting the manuscript, reviewing and editing the draft manuscript, submitting the final manuscript to the journal, and responding to reviewers' comments.

In paper 2, the study sought to assess the main risk factors of non-utilisation of postnatal care in DRC, Kenya and Tanzania using the Decision Tree.

2. Mlandu C, Matsena-Zingoni Z, Musenge E (2023). Predicting risk factors of non-utilisation of postnatal care in three sub-Saharan Africa countries: Application of the Decision Tree. (Under Review-*BMJ Public Health*).

PhD student: Responsible for the conceptualization of the manuscript which included defining the objectives and methodology to be used, performing data management (including data cleaning, coding, and manipulation), data analysis (including Machine Learning), drafting the manuscript, reviewing and editing the draft manuscript and submitting the final manuscript to the journal.

In paper 3, the study sought to predict the drop out from the maternal, newborn and child healthcare continuum in DRC, Kenya and Tanzania using various Machine Learning models.

3. Mlandu, C., Matsena-Zingoni, Z. & Musenge, E. Predicting the drop out from the maternal, newborn and child healthcare continuum in three East African Community countries: Application of machine Learning models. *BMC Med Inform Decis Mak* **23**, 191 (2023). https://doi.org/10.1186/s12911-023-02305-1.

PhD student: Responsible for the manuscript's conceptualisation which included defining the objectives and methodology to be used, conducting data management (including data cleaning, coding, and manipulation), data analysis (including multilevel logistic regression and Machine Learning), drafting the manuscript, reviewing, and editing the draft manuscript, submitting the final manuscript to the journal, and responding to reviewers' comments.

In paper 4, the study sought to describe the mediation role of maternal, newborn and child healthcare services utilisation on neonatal mortality in the DRC, Kenya and Tanzania.

4. Mlandu C, Matsena-Zingoni Z, Musenge E (2023). Mediation role of maternal and newborn and child healthcare services utilisation on neonatal mortality in three sub-Saharan African countries: A generalized structural equation modeling approach. *(Under Review- Maternal Health, Neonatology and Perinatology)*

PhD student: Responsible for the manuscript's conceptualisation which included defining the objectives and methodology to be used, conducting data management (including data cleaning, coding, and manipulation), data analysis (including generalised structural equation modeling), drafting the manuscript, reviewing and editing the draft manuscript and submitting the final manuscript to the journal.

ABSTRACT

Background:

The risk of child deaths within the first month of life is elevated than the later stages of childhood. Globally, Sub-Saharan Africa (SSA) has the highest neonatal mortality. Majority of the countries in SSA including the DRC, Kenya and Tanzania are struggling to meet Sustainable Development Goal (SDG) 3.2 of reducing the neonatal mortality rate to 12 deaths per 1,000 live births by 2030 (2). Most causes of neonatal deaths are preventable and treatable. Universal coverage, timely and effective utilisation of maternal, newborn, and child healthcare (MNCH) services during pregnancy, delivery, and postpartum has the potential to save many lives of newborns in high-burden countries.

Antenatal care (ANC) is the first service offered to pregnant women in MNCH. The timing and frequency of ANC visits is critical for the mother and her unborn child. The WHO recommends that women initiate ANC within 16 weeks of pregnancy and attend a minimum of four ANC visits for timely and optimum care before delivery (3, 4). The WHO also recommends that pregnant women receive assistance from a skilled worker during delivery and get postnatal checks with their newborns within 6 weeks of delivery (5, 6). Furthermore, utilising the Continuum of Care (CoC) for MNCH could significantly reduce maternal and newborn deaths in SSA. In the context of MNCH, the CoC is an approach that ensures continuous care from the period of pregnancy, through to childbirth, postnatal period, infancy, and the childhood period (7).

Despite the recognition of the use of vital services in MNCH, timely and adequate uptake of MNCH services remains poor and the coverage of MNCH is far from universal in SSA. Most pregnant women initiate ANC after 16 weeks and hence fail to receive timely ANC interventions (8). Uptake of ANC visits, skilled birth attendance (SBA) and postnatal care (PNC) is suboptimal (8-11).

Studies in SSA have explored various factors associated with MNCH services utilisation, however, our understanding of MNCH services utilisation in SSA is still limited. Trends in utilisation of MNCH services over time such as late ANC uptake have not been thoroughly assessed. Late uptake of ANC is still a common problem in SSA. Tracking women’s progress in the timing of ANC will ascertain if they are any changes in women’s late uptake of ANC and the contributing factors. This information will guide future policies and programmes which focus on improving the timely uptake of ANC in the SGD era. There is also a dearth of empirical evidence on the factors associated with the utilisation of ANC, skilled delivery and postnatal care in the CoC using nationally representative data. The CoC views both the mother and child as a collective rather than as separate/ individual entities. Understanding factors that contribute to the full utilisation of drop out from the CoC is essential for the formulation of interventions than enhance the CoC.

Furthermore, studies which investigated either the individual utilisation of MNCH services such as timing of ANC, ANC visits, SBA and PNC services or the CoC have tended to use more of the traditional analysis methods such as the logistic regression. The application of more versatile analysis methods such as Machine Learning is not common. Machine Learning methods are capable of extracting information that commonly used methods (logistic regression) fail to do by uncovering hidden patterns and relationships, particularly in large data sets (12). The application of Machine Learning methods can offer opportunities of enhancing existing methods (conventional regression methods) for predicting and classifying MNCH utilisation leading to more effective interventions to improve MNCH utilisation.

There is also a limited understanding on the interrelationships between MNCH services utilisation and neonatal outcomes. The associations between MNCH services utilisation and newborn outcomes such as neonatal mortality are commonly assessed using traditional approaches that assume direct associations. Specific analytical methods, such as Generalised Structural Equation Modeling (GSEM) can be used to model complex relationships such as interrelated links between utilisation of different MNCH services and neonatal outcomes. GSEM gives a clear understanding of how different services of MNCH are related to one another with neonatal outcomes by estimating both direct and indirect paths associations for more effective targeted interventions. Given the critical role of MNCH in ending preventable neonatal mortality, the overarching aim of this study was to describe the utilisation of MNCH services and their associations with neonatal mortality using GSEM and Machine Learning models in three sub-Saharan African countries: the DRC, Kenya, and Tanzania.

Methods:

The study utilised cross-sectional secondary data of reproductive-age women from the Democratic Republic of Congo (DRC) (2007-2013/14), Kenya (2008-2014) and Tanzania (2010-2015/16) Demographic Health Surveys.

Firstly, the multivariate logistic regression analysed factors associated with late ANC initiation accounting for clusters, survey weights and stratification for the different rounds of the Demographic Health Surveys. Trends in late initiation of ANC over time in each country were assessed by comparing the earlier and later surveys using differences in prediction scores (prediction probabilities generated after running the multivariate logistic regression models).

Secondly, the study assessed the main predictors of non-utilisation of PNC using the Decision Tree. The model performance of the Decision Tree was compared to the Logistic Regression using Accuracy, Sensitivity, Specificity and area under the Receiver Operating Characteristics.

Thirdly, factors associated with the drop out from the MNCH continuum, defined as not fully utilising either ANC, SBA, or PNC services, were analysed using multivariate logistic regression accounting for clusters, survey weights and stratification. Machine Learning analysis was used to predict the drop out from the MNCH continuum using features (predictors) that were found significant in the multivariate logistic regression. Five classification Machine Learning models were built and developed including the Artificial Neural Network, Decision Tree, Logistic Regression, Random Forest and Support Vector Machine to predict the drop out from the MNCH continuum. The prediction accuracies of the models were then compared using parameters including Accuracy, Precision, Recall, Specificity, F1 score and area under the Receiver Operating Characteristics.

Fourthly, the Generalised Structural Equation Modeling (GSEM) was used to assess the mediatory role of MNCH services utilisation on neonatal mortality. The endogenous variables were ANC attendance, SBA and PNC attendance, low birth weight and neonatal mortality. The GSEM analysis also accounted for survey weights and considered cluster random effects.

Results:

The findings showed a reduction in late ANC initiation (67.8%-60.5%) between 2008-2014 in Kenya as well as in Tanzania (60.9%-49.8%) between 2010-2016, but an increase was observed in the DRC (56.8%-61.0%) between 2007-2014. In the DRC, higher birth order was associated with ANC initiation delays from 2007-2014, whilst rural residency, lower maternal education and household income was linked to ANC initiation delays in 2014. In Kenya, lower maternal education and household income was associated with ANC initiation delays from 2008-2014, whilst rural residency and increased birth order were linked to ANC initiation delays in 2014. In Tanzania, higher birth order and larger households were linked to ANC delays from 2010-2016, whilst ANC initiation delays were associated with lower maternal education in 2010 and lower-income households in 2016.

The results also showed that the Decision Tree models had higher prediction accuracy of non-utilisation of PNC than the Logistic Regression models. Using the Decision Tree, low quality of ANC, home deliveries and unemployment were associated with the highest probability of not utilising PNC (92.0%) in the DRC. In Kenya, home deliveries, unemployment and lack of access to mass media were associated the highest likelihood of not utilising PNC (87.0%). In Tanzania, home deliveries, low quality of ANC and unwanted pregnancies exhibited the highest likelihood of not utilising PNC (100.0%).

The results also revealed very high rates of dropping out from the MNCH continuum in the DRC (91.0%), in Kenya (72.3%) and Tanzania (93.7%). Rural residence, lower maternal education and non-exposure to mass media were common predictors of dropping out from the MNCH continuum across the three countries. Further, the influence of factors such as

household wealth, household size, access to money for medication, travel distance to health facilities, and parity and maternal age varied by country. Results from the Machine Learning analysis showed that the Logistic Regression had the least prediction accuracy, while the Random Forest exhibited the highest prediction accuracy. Using the Random Forest, the study further ranked the most important predictors of the drop out from the MNCH continuum. Household wealth, place of residence, maternal education and exposure to mass media were the top four most important predictors.

The results also showed direct and indirect associations between MNCH services utilisation and neonatal mortality. ANC attendance mediated the total effects of PNC attendance on neonatal mortality by 8.8% in Kenya and 5.5% in Tanzania. ANC attendance and SBA also sequentially mediated the total effects of PNC attendance on neonatal mortality by 1.9% in Kenya and 1.0% in Tanzania. The results in Tanzania also showed ANC attendance mediated 2.8% of the total effects of LBW on neonatal mortality. No presence of mediation was observed in the DRC; however, ANC attendance moderated the relationship between parity and neonatal mortality.

Conclusions:

The study found that late uptake of ANC decreased between the two survey rounds in Kenya and Tanzania but increased in the DRC. Women from various geographic, educational, parity, and economic groups showed varying levels of late ANC uptake. Increasing women's access to information platforms and strengthening initiatives that enhance female education, household incomes, and localise services may enhance early ANC uptake.

The Decision Tree models showed higher prediction accuracy of non-utilisation of PNC than the Logistic Regression models in the DRC, Kenya and Tanzania. Using the Decision Tree, women who had poor quality of ANC, home deliveries, unemployment, unplanned pregnancies, and no mass media access were identified as high-risk subpopulations of non-

utilisation of PNC. Improving access and quality of care, incorporation of TBAs into the formal health systems, government health financing, increasing access to mass media and integrating maternal healthcare services with family planning services should be considered as top priority interventions to improving the utilisation of PNC.

Most women and children drop out of the MNCH continuum in the DRC, Kenya and Tanzania. Rural residence, lower maternal education and non-exposure to mass media were common factors linked to the high dropout in the MNCH continuum. The use of Machine Learning can help support evidence-based decisions in MNCH interventions. Rapid response mechanisms such as web-based applications can also be developed through the use Machine Learning whereby a pregnant woman's future utilisation of the services in CoC is assessed and monitored in real-time.

The GSEM findings showed interconnections between MNCH services utilisation such as timing of ANC, ANC visits, SBA, PNC and neonatal mortality. This suggests that more than direct and indirect factors are accountable for the associations between MNCH services utilisation and neonatal mortality. The mediation role of MNCH services on neonatal mortality indicates critical areas for targeted interventions to reduce neonatal mortality.

Overall, the study aimed to describe the utilisation of the MNCH services and associations with neonatal mortality in the DRC, Kenya and Tanzania. The study showed declines in late ANC uptake in two countries, however, early uptake of ANC is far is still not universal. The study also showed very low levels of retention in the CoC, and most women and children drop out in the CoC at postpartum period.

The findings also showed the existence of social, health system and individual inequalities in MNCH and their impact on early childhood survival. Women who are vulnerable to unequal and poor MNCH services utilisation are characterised by poverty, rural residence, long travel distances to health facilities, unaffordable medical expenses, home deliveries, low quality of

care, low education, high parity, younger age, unemployment, limited exposure to mass media, and unplanned pregnancies. Context-specific intervention programs such as female education, government health financing, MNCH promotion programs through mass media and improved accessibility and quality of care in health facilities, particularly for the most vulnerable groups of the populations such as women of low socioeconomic status and women from underserved rural areas are essential to improve the overall health of mothers and children and meeting the SDG-3 goals.

Modern biostatistical models like Machine Learning provide essential tools to understand public health problems. These techniques should be applied to complement the conventional statistical methods, particularly the tree-based models like the Decision Tree and Random Forest for predicting and classifying the utilisation of MNCH services. The GSEM established interconnections between timing of ANC, ANC visits, SBA and PNC and neonatal mortality. The timing of the first ANC contact is an important starting point to a continuation through the COC. It makes women better informed about pregnancy and the subsequent use of MNCH services. All stakeholders should work more on promoting early uptake of ANC by setting up initiatives that increase women's access to information platforms, enhance female education, improve household incomes, and bring services closer to communities.

Keywords: Generalised Structural Equation Modeling, Machine Learning, Maternal, Newborn and Child Healthcare

ACKNOWLEDGEMENTS

This has been a challenging but rewarding experience. I would like to express my gratitude to my supervisors, colleagues, family, and friends who have supported me throughout this journey. I am especially grateful to the individuals listed below, without whose assistance, guidance, counsel, and encouragement this would not have been possible. I appreciate Professor Eustasius Musenge's willingness to serve as my primary advisor. His continued support as a supervisor and mentor throughout the doctorate is very much appreciated. He constantly challenged me to explore new ideas and go beyond my comfort zone. I would also like to appreciate Dr Zvifadzo Matsena-Zingoni for her tremendous support and guidance throughout my PhD studies. She consistently provided timely, meticulous contributions to my work, which significantly enhanced its output. I am also appreciative of the support provided by the PhD administrator (Paul Bohloko).

I am thankful to the SAMRC Bongani Mayosi National Scholar program for the fellowship that provided financial and educational support for my doctoral studies. I am also grateful to the Postgraduate Financial Aid Office at the University of Witwatersrand for partially funding my PhD studies. I am also grateful to MEASURE DHS for providing the datasets for this study. Lastly, l am also thankful to my family: my mother, Joyce Mlandu, for your love and encouragement; my brother, Audwell Mlandu; my sister, Mildrate Garura; and my friend, Dr Blessing Kaunda, for your words of encouragement and support.

LIST OF ABBREVIATIONS

AIC:Akaike Information Criterion

AIDS:Acquired Immuno-Deficiency Syndrome

ANC:Antenatal Care

AOR: Adjusted Odds Ratio

AUROC:area under the Receiver Operating Characteristics

BIC:Bayesian Information Criterion

CHWs: Community Health Workers

CoC:Continuum of Care

DRC:Democratic Republic of Congo

DHS:Demographic and Health Survey

EAs:Enumeration Areas

EACs:East African Community countries

GSEM:Generalised Structural Equation Modeling

HCPs: Healthcare Professionals

HICs:High Income Countries

HIV:Human Immuno-Deficiency Virus

ICC:Intra-class Correlation Coefficient

IRB: Institutional Review Board

LMICs:Low-and-Middle-Income Countries

MDGs:Millenium Development Goals

MMR:Maternal Mortality Ratio

MNCH:Maternal, Newborn and Child Healthcare

ML:Machine Learning

NNR:Neonatal Mortality Rate

OOP:Out-of-Pocket

PEs: Processing Elements

PNC:Postnatal Care

SBA:Skilled Birth Attendance

SDGs:Sustainable Development Goals

SMOTE:Synthetic Minority Oversampling Technique

SSA:sub-Saharan Africa

ROSE: Random Over-Sampling Examples

TBAs: Traditional Birth Attendants

UNICEF:United Nations Children's Fund

USAID:United States Agency for International Development

WHO:World Health Organisation

DEFINITION OF TERMS

ANC: care received by the mother at onset and during pregnancy

CoC: refers to the continuity of care throughout the lifecycle, including pregnancy, childbirth, the postnatal period, and childhood

Drop out from the MNCH continuum: woman/child who does not fully utilise either ANC, SBA or PNC services.

Gestational age: a measure of the age of a pregnancy taken from the beginning of the woman's last menstrual period

KMeans SMOTE: an oversampling technique for class-imbalanced data. It aids classification by generating minority class samples in safe and crucial areas of the input space

Late ANC initiation: starting ANC at 4 or more months of pregnancy

LBW: birth weight which is <2,500g

Machine Learning: defined as the practice of extracting data, learning from it, and then predicting patterns using algorithms

Mediator: refers to a variable that explains the relationship between the outcome and explanatory variable

Moderator: refers to a third variable that affects the strength of a relationship between an outcome and an explanatory variable

PNC: care received by the mother or baby receiving care before or after six weeks of delivery

Random sampling: entails choosing samples at random from the minority class, replacing them, and adding them to the training dataset

SBA: receiving care from a doctor, nurse or midwife during labour

Universal Health Coverage: means that all people have access to the full range of quality health services they need, when and where they need them, without financial hardship

TABLE OF CONTENTS

CANDIDATE DECLARATION ii

DEDICATION iii

PREFACE iv

My contribution to the THESIS v

ABSTRACT vi

ACKNOWLEDGEMENTS xiv

LIST OF TABLES xxi

LIST OF FIGURES xxi

CHAPTER 1: BACKGROUND 2

1.0 Introduction 2

CHAPTER 2: LITERATURE REVIEW 5

2.0 Burden of neonatal mortality 5

2.1 Causes of neonatal mortality 5

2.2 MNCH services 6

2.2.1 Antenatal care 7

2.2.2 Skilled Birth Attendance 8

2.2.3 Postnatal Care 9

2.2.4 Continuum of Care for MNCH 9

2.3 Statistical Methods 10

2.3.1 The Logistic regression model 10

2.3.2 Machine Learning 10

2.3.3 Generalised Structural Equation Modeling 13

2.4 Problem Statement 14

2.5 Justification 16

2.6 Overall aim 19

2.6.1 Specific Objectives 19

2.7 The conceptual framework describing the utilisation of MNCH services and their associations with neonatal mortality. 20

CHAPTER 3: METHODOLOGY 21

3.0 Study setting 21

3.1 Study design and data source 23

3.2 Data sampling and collection 24

3.3 Data management and analysis 24

3.3.1 Objective 1 24

3.3.2 Objective 2 26

3.3.3 Objective 3 29

3.3.4 Objective 4 32

3.6 Papers and Supplementary Results 37

CHAPTER 4: EMPIRICAL EVIDENCE for Paper/objective 1 38

4.0 Introduction 39
4.1 Materials and Methods 41
4.1.1 Study design and setting 41
4.1.2 Inclusion and exclusion criteria 41
4.1.3 Data Source and Sampling 41
4.1.4 Study Variables 42
4.1.5 Data analysis 42
4.2 Results 43
4.2.1 Descriptive analysis 43
4.2.2 Determinants of late ANC initiation in the combined countries 51
4.3 Discussion 56
4.3.1 Strengths and Limitations of the Study 59
4.4 Conclusions 59
4.5 Chapter Summary 60
CHAPTER 5: EMPIRICAL EVIDENCE for paper/objective 2 61
5.0 Introduction 62
5.1 Methods and Materials 64
5.1.1 Study design and Sampling procedure 64
5.1.2 Inclusion and exclusion criteria 65
5.1.3 Study variables 65
Outcome variable 65
Explanatory variables 65
5.1.4 Data management and analysis 66
5.2 Results 69
5.2.1 Demographic and Socioeconomic Characteristics of Study Participants 69
5.2.2 Health service-related Characteristics of Study Participants 71
5.2 Decision Tree Analysis 72
5.4 DISCUSSION 77
5.5 Conclusions 82
5.6 Chapter Summary 83
CHAPTER 6-EMPIRICAL EVIDENCE for objective 3 84
6.0 Introduction 85
6.1 Methods 87
6.1.1 Study Design and Settings 87
6.1.2 Study Population 87
6.1.3 Data Source and Sampling 88
6.1.4 Measurement of Variables 88
6.1.5 Statistical analysis 89

6.2 Results 90
6.2.1 Characteristics of the women 90
6.2.2 Patterns of the drop out from the MNCH continuum 92
6.2.3 Predictive modelling 96
6.3 Discussion 98
6.3.1 Strengths and Limitations 101
6.4 Conclusions 102
6.5 Chapter Summary 103
CHAPTER 7-EMPIRICAL EVIDENCE for objective 4 104
7.0 Introduction 106
7.1 Methods 108
7.1.1 Study area and setting 108
7.1.2 Data source design and sampling 108
7.1.3 Inclusion and exclusion criteria 108
7.1.4 Measurement and variables 109
Outcome and explanatory variables 109
Exogenous variables 109
7.1.5 Data analysis 110
7.2 Results 112
7.2.1 Characteristics of study participants 112
7.2.3 Mediation analysis 127
7.2.4 Comparison of GSEM models with or without cluster random effects 132
7.3 Discussion 132
7.3.1 Strengths and Limitations 135
7.4 Conclusions 135
7.5 Chapter Summary 136
CHAPTER 8: GENERAL DISCUSSION 137
8.1 Summary of findings 137
8.1.1 Low utilisation of MNCH services 137
8.1.2 Emerging themes on MNCH services utilisation 139
8.1.2.2 Organisational or health system factors 142
8.1.2.3 Individual factors 146
8.1.3 Linkages of MNCH services utilisation 148
8.1.4 Comparison of the Machine Learning and Logistic regression 149
8.1.5 Application of the GSEM 152
8.2 Recommendations for Future research projects 152
8.3 Strengths and Limitations 153
8.4 Conclusions 154

APPENDIX 156
REFERENCES 168

LIST OF TABLES

Table 3.1: DHS surveys data used in the study analysis 23
Table 4.2: Late ANC initiation by background characteristics in the DRC, Kenya and Tanzania 49
Table 4.3: Determinants of late ANC initiation in the combined countries 52
Table 4.4: Determinants of late ANC initiation in the DRC, Kenya and Tanzania 54
Table 5.1: Demographic and Socioeconomic characteristics of Study Participants 70
Table 5.2: Health service-related characteristics of Study Participants 71
Table 5.2: Performance parameters of the Decision Tree and Logistic Regression models in the DRC, Kenya, and Tanzania 76
Table 6.1: Characteristics of women in countries under study 91
Table 6.2: Factors associated with the drop out from the MNCH continuum in the DRC, Kenya and Tanzania. 94
Table 6.3: Prediction analysis of the drop out from the MNCH continuum using combined DHS surveys data for the DRC, Kenya and Tanzania (2013-2016) 96
Table 7.1: Characteristics of study participants in the DRC, Kenya and Tanzania 114
Table 7.2: Analysis of the interrelationships between MNCH services utilisation and neonatal mortality using GSEM in the DRC 117
Table 7.3: Analysis of the interrelationships between MNCH services utilisation and neonatal mortality using GSEM in Kenya 121
Table 7.4: Analysis of interrelationships between MNCH services utilisation and neonatal mortality using GSEM in Tanzania 124
Table 7.5: Neonatal mortality mediation analysis using the GSEM in the DRC, Kenya and Tanzania 131
Table 7.6: Comparison of GSEM models with or without cluster random effects 132
Table 8.1: Consolidated Study Findings 138

LIST OF FIGURES

Figure 2.1: Conceptual framework describing the utilisation of MNCH services and their associations with neonatal mortality, adapted from McLeroy's Socio-Ecological Model 21
Table 3.1: DHS surveys data used in the study analysis 23
Figure 3.1: A conceptual framework for the mediation role of MNCH services utilisation on neonatal mortality [adapted from the Socio-Ecological Model (1)]. 34
Figure 3.2a-b: a) Single Mediation b) Sequential Mediation 36
Figure 4.1.: Map showing the prevalence of late ANC initiation in the DRC, Kenya, and Tanzania in the first survey in this study: Base map source at https://spatialdata.dhsprogram.com. 43
Figure 4.2: Map showing the prevalence of late ANC initiation in the DRC, Kenya, and Tanzania second survey in this study. Base map source at https://spatialdata.dhsprogram.com. 44
Figure 4.3: Trends of late ANC initiation in the DRC, Kenya, and Tanzania between the first and second surveys. 45
Figure 5.1: Decision Tree for predicting the risk of non-utilisation of PNC among women in the DRC; 1) root, 2) low quality ANC, 3) high quality ANC, 4) home delivery, 5) facility delivery, 6) rural residence, 7) urban residence, 8) no , 9) had employment, 10) normal delivery, 11) caesarean section, 12) home delivery, 13) facility delivery, 14) primary or no education, 15) secondary or higher education. 73
Figure 5.2.Decision Tree for predicting the risk of non-utilisation of PNC among women in Kenya; 1) root, 2) home delivery, 3) facility delivery, 4) unemployment, 5) employed 6) normal delivery, 7) caesarean

section, 8) non exposure to mass media, 9) exposure to mass media, 10) no SBA, 11) had SBA, 12) big problem seeking medical help alone, 13) no big problem seeking medical help alone, 14) no SBA, 15) had SBA. 74

Figure 5.3: Decision Tree for predicting the risk of non-utilisation of PNC among women in Tanzania; 1) root, 2) home delivery, 3) facility delivery, 4) low quality of ANC, 5) high quality of ANC 6) low quality of ANC , 7) high quality of ANC, 8) wanted pregnancy no more, 9) wanted pregnancy then, 10) big problem with seeking medical help alone 11) no big problem with seeking medical help alone, 12) big problem with medical costs, 13) no big problem with medical costs, 14) big problem with seeking medical help alone, 15) no big problem with seeking medical help alone. 75

Figure 6.1: Patterns of the drop out from the MNCH continuum among reproductive-age women in the last 5 years preceding each country's DHS survey. 92

Figure 6.2: AUROC parameters for the five ML classification models 97

Figure 6.3: Feature (variable) ranking using the Random Forest 97

Figure 7.1: A conceptual framework for the mediation role of MNCH services utilisation on neonatal mortality [adapted from the Socio-Ecological Model (1)]. 110

Figure 7.2: GSEM model of mediating role of MNCH services utilisation on neonatal mortality in the DRC. Key: variable type- light blue box (mediator), rose box (neonatal mortalty), mint box (direct predictors of neonatal mortality), white box (covariates) and lavender box (cluster random variable) 128

Figure 7.3: GSEM model of mediating role of MNCH services utilisation on neonatal mortality in Kenya. Key: variable type- light blue box (mediator), rose box (neonatal mortality), white box (covariates), mint box (direct predictors of neonatal mortality), and lavender box (cluster random variable) 129

Figure 7.4: GSEM model of mediating role of MNCH services utilisation on neonatal mortality in Tanzania. Key: variable type- light blue box (mediator), rose box (neonatal mortality), white box (covariates), mint box (direct predictors of neonatal mortality) and lavender box (cluster random variable) 130

CHAPTER 1: BACKGROUND

The background gives a brief introduction of the broad context underlying the research, overview of the problem and research gap.

1.0 Introduction

The neonatal period, defined as the first 28 days of life, is the most vulnerable period of a child's life (13). Children face the highest risk of death in their first month of life than later stages (13). Globally 2.5 million children died in the first month of life in 2017, approximately 7000 every day (2). The chances of survival from birth varies widely by region. Sub-Saharan Africa (SSA) recorded the highest Neonatal Mortality Rate (NMR) in 2017 at 27 (25–32) deaths per 1000 live births, followed by Southern Asia with 26 deaths per 1,000 live births (2).

Current projections show that many countries in SSA like the DRC, Kenya and Tanzania will not meet the SGD-3 goal of reducing the NMR to as low as 12 deaths per 1,000 live births by 2030 (2). The DRC, Kenya and Tanzania rank in the top six countries in SSA with the highest number of neonatal deaths (14). In Africa, the major causes of newborn deaths include sepsis/pneumonia, diarrhoea, tetanus, preterm birth, asphyxia and congenital defects (15, 16). Most newborns deaths result from preventable and treatable causes (17).

Timely and adequate utilisation of MNCH is critical for enhancing maternal and neonatal survival (17, 18). ANC is an excellent entry point for healthcare professionals to offer a variety of interventions to expectant mothers (19). Early ANC initiation within 16 weeks of pregnancy assists health providers in giving timely ANC interventions (3, 20). The WHO also recommends that all pregnant women should attend a minimum of four ANC visits (4, 21) because it affords expectant mothers adequate time to be examined and cared for before delivery. Chances of a safe and normal delivery are also higher when expectant mothers are assisted by qualified and experienced professionals such as doctors, nurses, or midwives (5). The mother and child should also continue to receive care within six weeks after delivery to prevent and manage complication during the postpartum period (22). Furthermore, uptake of antenatal, delivery and postnatal care in a CoC could accelerate the achievement of SDG 3.1 of reducing maternal deaths by 70 deaths per 100 000 live births and SDG 3.2 of ending preventable neonatal mortality by 2030, and reducing neonatal mortality to 12 deaths per 1,000 live births (7, 23).

It is vital that governments increase the access and provision of high quality MNCH services, including antenatal, delivery, and postnatal services (24). However, this remains a challenge despite the commitments encapsulated in the SDG goals (25) and country-level policies in SSA (24). Although these initiatives are designed to increase the utilisation of MNCH services and ensure universal coverage, the timely and adequate utilisation of MNCH services remains poor in SSA. Reports from literature show that over 60% of the pregnant women presented late for their first ANC visit (8), over 40% of the women did not attend a minimum of four ANC visits (26) and over 35% of the women delivered without the assistance from a skilled worker (27) and more than 45% the women did not receive postnatal checkups (11).

Studies in different settings have investigated the factors associated with MNCH services utilisation including the timing of ANC, number of ANC visits, skilled delivery and postnatal care (8, 9, 28-30), however, our understanding of MNCH utilisation in SSA is still limited. Trends in utilisation of MNCH services such as late ANC uptake have not been thoroughly assessed. to monitor how women's late uptake of ANC has progressed (8). Moreover, few studies have investigated the utilisation of ANC visits, skilled delivery and postnatal care in the CoC to better understand and improve the utilisation of MNCH continuum (31-33). In addition, the use of more versatile analytical approaches like Machine Learning methods in predicting and classifying the utilisation of MNCH services is less common (34, 35).

Furthermore, associations between MNCH services utilisation and neonatal outcomes such as neonatal mortality have often been assessed using traditional approaches (regression models) that assume direct associations (36, 37). However, few studies have assessed interrelated (direct and indirect) pathways between MNCH services utilisation and neonatal outcomes (38, 39). Analytical approaches like GSEM can capture complex and dynamic relationships such as the links between MNCH services utilisation and neonatal outcomes. Unlike regression models, GSEM can estimate both direct, indirect and total path effects (40) which gives a clear understanding of how different services of MNCH are related to one another with neonatal outcomes for more effective targeted interventions (40).

Given the importance of MNCH in ending preventable neonatal mortality and accelerating the achievement of SDG-3 goals, the study sought to describe the utilisation of MNCH services and their associations with neonatal mortality using GSEM and Machine Learning models in three sub-Saharan African countries namely the DRC, Kenya, and Tanzania.

CHAPTER 2: LITERATURE REVIEW

This review of the literature describes the burden and causes of neonatal mortality. It describes MNCH services utilisation and associations with neonatal outcomes. It also describes the statistical methods used in assessing utilisation of MNCH services. It also gives more detail on the problem statement, justification of the study and outlines the aim and objectives of the PhD study, and conceptual framework.

2.0 Burden of neonatal mortality

Neonatal deaths have generally declined over time across the globe. From 1990 to 2017, global NMR decreased from 37 (36-38) deaths per 1,000 live births in 1990 to 18 (17-20) in 2017 (2). However, the chances of survival for newborns continue to significantly vary by region. In 2017, the NMR (27 deaths per 1000 live births) in SSA was 9 times higher compared to HICs (3 deaths per 1000 live births) (2). Improving the chances of newborns' survival is still a challenge in SSA. Most countries in SSA are falling short of the SGD-3 target of reducing neonatal mortality to 12 per 1,000 live births by 2030 (2). In 2017, countries like the DRC (29 deaths per 1,000 live births), Kenya (21 deaths per 1,000 live births) and Tanzania (21 deaths per 1,000 live births) recorded NMRs way above the SDG-3 target (2). The DRC, Kenya and Tanzania are ranked among the top six countries in Africa with the highest number of neonatal deaths (14). Without intensified goals, most countries in SSA will not meet the SDG-3 goals (2, 14, 18, 40).

2.1 Causes of neonatal mortality

Globally, the main causes of neonatal deaths globally are prematurity (28%), sepsis (26%)and asphyxia/birth related hypoxia (23%)(17). In Africa, 1,12 million newborn deaths occur annually (41). The leading causes of neonatal deaths include sepsis/pneumonia, diarrhoea, tetanus, preterm birth, asphyxia and congenital defects (15, 16). These causes account for approximately 80% of neonatal deaths (41).

Most causes of neonatal deaths are preventable and treatable (17). Pneumonia and diarrhoea are diseases that are prone in children with weakened immune sysytems (42, 43). Poor sanitation, malnutrition, low birth weight (LBW) and insufficient immunisation increase children's susceptibility to pneumonia and diarrhoea (42, 43). Tetanus is a fatal infectious disease which is preventable through vaccination of pregnant mothers and children (44). Prematurity is an underlying cause of 50% of neonatal deaths, very premature neonates (Gestational age <32 weeks) have a particularly high risk of death (45). Prematurity for a variety of reasons such as maternal chronic conditions and infections and pregnancy complications (45). Birth asphyxia, defined as the failure to establish breathing at birth**,** is one of the primary causes of early neonatal mortality. The most common cause of perinatal asphyxia is complications during childbirth (46). Congenital abnormalities are caused by problems during the fetus's development before birth due to pregnancy complications and environmental factors (47).

2.2 MNCH services

The "Health Goal" or SDG-3 aims toward "ensuring healthy lives and promoting well-being for all at all ages" (25). The centerpiece of SDG-3 is "Universal Health Coverage" (UHC), which signifies that all people should have access to quality health services, when in need, without facing any financial hardship. It covers the full continuum of essential health services, from health promotion to prevention, treatment, rehabilitation, and palliative care across the life course (25). Universal coverage of MNCH is an essential pre-requisite for women and children in the advancement of better health outcomes and achievement of SDG-3 goals of reducing maternal deaths and ending preventable neonatal mortality (25, 48). According to

WHO, MNCH interventions offered during antenatal, delivery and postnatal period could improve maternal and child outcomes. However, poor uptake of MNCH services, particularly in developing countries is linked to maternal and early childhood morbidity and mortality (17, 18, 49, 50).

2.2.1 Antenatal care

Antenatal care is the first package of MNCH services offered to pregnant women (7). It is routine care provided to pregnant mothers from conception to the start of labour (9). ANC aims to promote and protect the health of the expectant mother and her unborn child through various interventions (9) such as the identification of pre-existing health conditions, early detection of complications, health promotion and disease prevention, and birth preparedness and planning (21). Identification of pre-existing health conditions includes checking for weight, nutrition status, anemia, hypertension, syphilis, and HIV status (3). Early detection of complications arising during pregnancy includes screening for pre-eclampsia and gestational diabetes (3). Health promotion and disease prevention includes tetanus vaccine, prevention and treatment of malaria, nutrition counseling, micronutrient supplements and family planning and counselling (3). Birth preparedness and planning includes birth and emergency plans, breastfeeding counselling and antiretrovirals for HIV positive women and prevention of mother-to-child transmission of HIV (3).

The first timing of ANC refers to the first time a pregnant woman visits a prenatal clinic to receive care from health care professionals (20). It is used to determine the health state of the mother and fetus, to estimate gestational age and date of delivery, and to schedule for future ANC visits (20). The timely commencement of ANC is critical for early detection of pregnancy-related issues and unfavourable pregnancy outcomes such as LBW, stillbirth, and intrauterine fetal death (20). The WHO recommends that pregnant women begin ANC early, within the first 16 weeks of pregnancy, for the optimum health outcomes (3, 20). However, most women begin ANC late in SSA. Reports from a study conducted in SSA indicate that over 60% of the pregnant women present late for their first ANC visit (8). Late ANC uptake has been associated with characteristics such as low education, rural residence, low income, multiparity, unemployment, unplanned pregnancies, and travel expenses (8, 51).

The WHO also recommends a minimum of four ANC visits for low risk pregnancies (4, 21). This allows for adequate monitoring of the mother's health and fetus up to the time of delivery (4, 21). However, ANC utilisation is low in SSA (9, 52-54). Findings from an earlier study found that only 55.5% of the women attended a minimum of four ANC visits (26). Attending fewer ANC visits has the potential to increase maternal and child morbidity and mortality (55, 56). Findings from a cohort study in Ethiopia showed attendance of four or more ANC visits was significantly associated with a reduction in postpartum haemorrhage, preterm birth, and LBW and neonatal mortality (57). Reasons cited for low attendance of ANC visits in literature include late ANC initiation, poor income, being single, low education, living in the rural areas, older age, higher parity, living in larger households, lack of health insurance cover, low quality of ANC and longer travel distance to the health facility (9, 19, 54, 58).

2.2.2 Skilled Birth Attendance

Skilled birth attendance is recommended during labour (59). Skilled workers such as doctors, nurses and midwifes are trained in managing normal (uncomplicated) pregnancies, childbirth and the immediate postnatal period, and in the identification, management and referral of complications in women and newborns (5). Skilled birth attendance has the potential to reduce maternal mortality and ensure child survival (15), as outlined in the United Nations' Sustainable Development Goals 3.1 and 3.2 ((SDG 3.1 and 3.2) (5, 25). However, reports from literature show that over a third of mothers in SSA deliver without the assistance from a skilled worker

(27). Barriers affecting lack of SBA include low ANC visits, lack of access to mass media, longer travel distances to the health facility, low quality of ANC, higher parity, medical expenses, low education, rural residence, cultural norms and low income (58-60).

2.2.3 Postnatal Care

Postnatal care is provided to the mother and newborns during the postnatal period, commencing immediately after childbirth and continuing for up to six weeks (42 days) after delivery (61, 62). It includes risk identification, preventive measures, health education and promotion, and complications management or referral. It reduces maternal and child morbidity and mortality while also improving overall health and well-being (62). Despite the aforementioned benefits, only 52.5% of women and newborns in SSA receive a postnatal checkup, and postnatal care is the least utilised MNCH service (11). Factors that contribute to low uptake of PNC found in literature include having informal education, low ANC visits, lack of SBA, home delivery, low quality of ANC, medical expenses, low autonomy in healthcare decision making, rural residence, higher parity, female headed-households and unemployment (11, 22, 58, 63).

2.2.4 Continuum of Care for MNCH

Improving the utilisation of the CoC for MNCH is one of the strategies and recommendations for achieving the global goal of reducing maternal mortality (70 maternal deaths per 100,000 live births) and ending preventable neonatal mortality and under-five mortality by 2030 (25). However, over 85% of the women and children in SSA do not fully utilise the CoC or the three services combined i.e. ANC, SBA and PNC (52, 64). The CoC is currently being given more attention and deemed advantageous over receiving each care provided in the MNCH separately (52, 64). The CoC for MNCH places focus on two critical dimensions: time and place, or level of care (65). The CoC for MNCH recognizes that there is a close inter-relationship between MNCH at different time periods and location (65). The time dimension emphasises the need of linking MNCH service packages over time during pregnancy, labor, and the postpartum period (65). The place or level of care dimension links various levels of services provided at home, communities and health facilities (65). Considering that the mother’s health affects the health of the child, and the importance of continuity in care, it is essential for their interventions to be enhanced through a CoC approach (64, 66).

2.3 Statistical Methods

Several statistical techniques are used in MNCH research to quantify relationships including the logistic regression models, Machine Learning algorithms and Generalised Structural Equation Modelling (GSEM). The application of more complex analytical methods such as Machine Learning and GSEM is becoming increasingly relevant in the analysis of healthcare data than common statistical models (logistic regression) (34, 67-69).

2.3.1 The Logistic regression model

The logistic regression model is a statistical analysis method used to analyse the association between a binary response outcome, and an explanatory variable. The logistic regression model is popular among health researchers (9, 70). It is based on the assumption that observations are independent (71). For the survey logistic regression model, one needs to account for sampling weights and the clustering of observations within survey Enumeration Areas (EAs) (71). Prediction scores (prediction probabilities) can also be generated after running the logistic regression model (72). The benefit of using prediction scores is that the predicted values of the outcome variable are adjusted for covariates or potential confounders (72).

2.3.2 Machine Learning

Machine Learning is a subset of artificial intelligence which is one of the fastest-growing technical fields (73). Machine Learning describes the automated process of identifying patterns in data to perform tasks, such as prediction and classification. It is effective in finding correlations in data, but not determining causation (68, 74). Machine Learning has become indispensable for managing large data sets (75). For some problems with small datasets, for instance, conventional techniques like the logistic regression provide a quick and cost-effective solution, whereas Machine Learning algorithms provide a better option for complex problems with large datasets and nonlinear interaction between different variables (35). Big data allows Machine Learning algorithms to uncover hidden patterns which stimulate the process of decision-making (76).

The data used in Machine Learning can be labelled or unlabelled. Labeled data contains meaningful tags or labels, while unlabeled data is simply raw data before labelling (76). Using this data the Machine Learning algorithm estimates a pattern regarding about the data and uses a function where the estimation is compared to the known answer i.e. the labeled data, to measure accuracy (76). The model then attempts to fit the estimation to known data points to increase accuracy (76). This is how the Machine Learning algorithm trains and produces models that assist the machine in mimicking human behaviour.

In today's environment, Machine Learning is being used in a range of fields (12, 76). Machine Learning methods mainly are for classification and prediction (e.g., medical diagnosis, forecasting pandemic, clustering analysis (e.g., securing the web by detecting unusual traffic, identification of cancer cells, partitioning customers) (12, 76) and natural language processing (e.g., speech recognition, language translation, sentiment analysis) (12, 76). However, few studies have employed Machine Learning in MNCH to predict or classify the utilisation of MNCH services (31, 32, 34).

Machine Learning can be categorised into several ways depending on how the system (ML model or agent) is trained (76). The main types of Machine Learning are supervised learning and unsupervised learning (76). In supervised learning, the datasets are labeled to train algorithms to classify data or predict outcomes, while in unsupervised learning models are not supervised using training datasets but instead, models learn from the hidden pattern and unknown information from the datasets (76).

Several complex Machine learning algorithms or models are used in supervised Machine Learning including the Decision Tree, Random Forest, Artificial Neural Network, Support Vector Machine (77). A Decision Tree is a supervised learning algorithm used for classification and prediction. It depicts a flowchart, starting with a root node that asks a specific question about the data. Based on the answer, the data is directed down different branches to subsequent internal nodes, which ask further questions and guide the data to subsequent branches (77, 78). This process goes on until the data reaches an end node, also known as a leaf node, where no further branching occurs. By asking a sequence of questions and following the corresponding branches, Decision Trees enable prediction and classification of outcomes based on the data's characteristics. Decision Tree algorithms can handle complex datasets with ease and simplicity. The simplicity and interpretability make Decision Trees essential for various applications in Machine Learning, particularly when handling complex datasets (77).

A Random Forest algorithm is an ensemble of Decision Trees used for prediction and classification (77). The Random Forest combines the predictions from multiple Decision Trees to make more accurate predictions. Numerous Decision Tree algorithms (sometimes hundreds or even thousands) are separately trained using different random samples from the training dataset (79). Once trained, the Random Forest takes the same data and inputs it into each Decision Tree. Each Decision Tree produces a prediction, and the Random Forest tallies the results. The most common prediction among all the Decision Trees is then selected as the final prediction for the dataset. Random Forests address overfitting that can occur with individual Decision Trees. Overfitting occurs when a Decision Tree becomes too closely aligned with its training data, making it less accurate when presented with new data (77).

An Artificial Neural Network is a Machine Learning algorithm used to classify and cluster data at high velocity (77). It makes decisions in a similar way to the human brain using processes that mimic the way biological neurons work together to identify phenomena, weigh options and come at conclusions. Each neural network consists of layers of nodes, or artificial neurons—an input layer, one or more hidden layers, and an output layer (80). Each node connects to others and has its own associated weight and threshold. If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network. Neural networks depend on training data to learn and improve their accuracy over time (77).

A Support Vector Machine is used for classification and predictive modeling tasks (77). Support Vector Machine algorithms work by creating a decision boundary called a "hyperplane." In two-dimensional space, this hyperplane is like a line that separates two sets of labeled data (77). The aim of the Support Vector Machine is to get the best possible decision boundary by maximizing the margin between the two sets of labeled data. The Support Vector Machine searches for the widest gap or space between the classes. Any new data point that falls on either side of this decision boundary is classified based on the labels in the training dataset (80). It's essential to note that hyperplanes can take on different shapes when plotted in three-dimensional space, allowing Support Vector Machines to handle more complex patterns and relationships in the data (77).

2.3.3 Generalised Structural Equation Modeling

Structural Equation Modeling (SEM) is a comprehensive statistical technique employed to estimate a network of relationships between observed (measured) variables and latent variables

(81). SEM has gained popularity in health sciences as a powerful multivariate analysis tool due to its ability to handle the investigation of both simple and complex causal models (38, 81). SEM is a progression beyond the well-established linear and logistic regression models. Unlike regression models, the SEM can estimate both direct, indirect, total path effects, quantify each factor's contribution to the covariance structure and identifies patterns of covariance among variables, explore potential latent confounders and can model all the equations simultaneously (81).

Generalized Structural Equation Modeling (GSEM) is an extension of SEM. SEM differs from GSEM in that the outcome variable in SEM is continuous and the regression model is linear, whereas GSEM is more general (81, 82). The response variables can be continuous, binary or ordinal. In addition, non-linear link functions are allowed. This refers to the ability to simultaneously handle nested or crossed group-level effects in a particular data set, i.e., latent and observed variables that may vary at various levels can be modelled simultaneously. Consequently, GSEM allows the incorporation of both unobserved and observed effects for subjects, subjects within groups, and groups within subgroups (82).

2.4 Problem Statement

Globally, 2.5 million newborn deaths occurred in 2017, approximately 7000 every day (2). The probability of survival from birth varies widely by region (2). Sub-Saharan Africa (SSA) had the highest NMR in 2017 at 27 (25–32) deaths per 1000 live births, followed by Southern Asia with 26 deaths per 1,000 live births (2). Also in SSA, nearly half of the newborn deaths occur in six countries namely the Nigeria, DRC, Ethiopia, Tanzania, Uganda and Kenya (14). The common causes of neonatal deaths in Africa include sepsis/pneumonia, tetanus, diarrhoea, preterm birth, asphyxia and congenital defects (15, 16). However, most of the leading causes of neonatal deaths are preventable and treatable (15, 17).

Using current projections, most countries in SSA will miss the SDG-3 target of reducing the NMR by 12 deaths per 1,000 live births (2). The DRC is one of the countries with the highest NMR in SSA, estimated at 29 deaths per 1000 in 2017, and 68 000 neonatal deaths Artificial annually (2). The adequate provision of MNCH has been adopted as one of the key strategies in preventing newborn deaths (83-85). The DRC's health sector is mainly financed by international aid, particularly in war torn areas (84). For instance, since 1963, the United Nations Children's Fund (UNICEF) has been providing health aid on maternal and child health to the DRC, particularly in impoverished areas and areas of conflict (84).

In Kenya, newborn deaths are a major public health problem (36). Although Kenya has reduced NMR from 33 deaths per 1,000 live births in 1990 to 21 deaths per 1,000 live births in 2017, about 32,000 children die in the first 28 days (2). The Kenyan government adopted several strategies to improve access of MNCH particularly among the poor such as provision of free maternal healthcare and national health schemes (86). In 2013, Kenya declared maternity services free of charge, in all public health institutions across the country to make maternal healthcare accessible to all women in the country (86). However, evaluations of the policy on the abolishment of maternal healthcare user fees uncovered several inadequacies. Women still grappled with other financial barriers like transportation costs and health facilities experienced problems with shortages of staff, drugs and supplies and ambulances for community-level emergencies (86, 87). This resulted in an increase in out-of-pocket (OOP) payments (88). Kenya has two primary health insurance schemes; the National Health Insurance Fund (NHIF), which was founded in 1966, and the National Social Security Fund (NSSF), which was founded in 1965 (88). Although several health insurance schemes are available, health insurance currently provides coverage for only 10% of Kenya's population.

In Tanzania, neonatal mortality is also of public health concern (89). Although the NMR decreased from 38 deaths per 1000 live births to 21 deaths per 1000 live births between 1990 and 2017, the reduction has been slow (2). The government of Tanzania has endeavoured to improve access of healthcare among mothers and children through initiatives such as the National Health Insurance Fund (mandatory for public officials) and Community Health Fund (optional for the informal sector) introduced in 2000, and the provision of free maternal and child healthcare in 2008 (90). However, low membership has been persistently observed in community health schemes, which may be attributable to poor quality of public healthcare, lack of provider choice, and inadequate reward packages (91). Moreover, although user fees have been eliminated, women in Tanzania continue to incur substantial OOPs expenses for medications, supplies, and the transportation of the mother and any birthing companions (90).

MNCH has been adopted as a key element in improving the lives of women and children in the SDG goals, however, timely and adequate utilisation of MNCH is still poor and far from universal in SSA. Reports from literature show that more than two-thirds of the women initiated ANC late (8) and only 56.7% of the women attended a minimum of four ANC visits (26), 63.9% received SBA (27) and only 52.5% of the women had postnatal checks after delivery (11). Moreover, over 85% of the women do not fully utilise all the services i.e.; ANC, SBA and PNC (52). Thus, it is conceivable the low utilisation of MNCH services is also contributing to SSA's ongoing problems of neonatal mortality.

2.5 Justification

While MNCH i.e. ANC, SBA and PNC has been highlighted as a key strategy in the SDG-3 goals to ending preventable neonatal mortality (25), uptake of MNCH services remains poor in SSA and is of public health concern (9, 28, 59, 62, 63). Although studies have been conducted to understand the utilisation of vital services in MNCH such as ANC, SBA and PNC

(9, 28, 62), our understanding of MNCH utilisation is still limited. The assessment of trends of MNCH services utilisation over time has not been extensively investigated in SSA countries, particularly the timing of the first ANC visit (92). Late ANC initiation is a common problem in SSA, and there are limited studies in SSA which have assessed trends of late ANC initiation and contributing factors (92). Using two rounds of DHS surveys, the study assessed the trends of late ANC initiation between the two rounds of DHS surveys and associated factors in the DRC, Kenya and Tanzania. This information will help policy makers in implementing interventions on the factors contributing to the observed trends of late ANC initiation and guide future policy in improving timely initiation of ANC in the SDG era.

The uptake of ANC visits, SBA and PNC remains suboptimal in SSA (30). While previous studies have identified various factors that influence the individual utilisation of ANC visits, SBA and PNC (9, 28, 62), few studies have assessed the drop out or full utilisation of the CoC in SSA (31, 32). The CoC is currently receiving more attention and is preferable to receiving each service provided in the MNCH separately i.e. ANC visits, SBA, and PNC (64, 66). Because the health of the mother is closely tied to that of the child, women stand to benefit more with their children when they receive optimal services from pregnancy, childbirth and the postnatal period. Given the importance and benefits of integrated utilisation of MNCH services, it is vital that the barriers preventing women/children from fully utilising the CoC are understood. Thus, this study sought to investigate the factors associated with the drop out of women/children in the CoC for MNCH.

Additionally, studies which assessed the utilisation of ANC visits, SBA and PNC services separately or the MNCH continuum have tended to use conventional analysis methods such as the Logistic regression (30-33, 60, 93). However, the use of more versatile analytical approaches for predicting and classifying MNCH utilisation using like Machine Learning are not common. Machine Learning algorithms like the Decision Tree, Random Forest, Support

Vector Machine and Artificial Neural Network have a larger capacity of extracting hidden patterns and relationships in data than conventional methods, particularly when utilising large datasets. Thus, the study sought to utilise Machine Learning in identifying the main factors of non-utilisation of PNC using the Decision Tree. The Decision Tree method was applied because of several advantages. The Decision Tree algorithm can identify high-risk population subgroups of women/children who do not utilise PNC. Interventions can be then developed according to the specific needs of each subgroup, particularly for vulnerable groups. The results from the Decision Tree are also simple and easy to interpret.

The study also applied several complex Machine Learning algorithms/models like the Decision Tree, Random Forest, Support Vector Machine and Artificial Neural Network to predict the utilisation of the CoC and use the best performing Machine Learning algorithm to rank the most important features/predictors. To the best of our knowledge, no study has used Machine Learning to predict the utilisation of the CoC for MNCH services in the three countries: DRC, Kenya and Tanzania. Application of Machine Learning will aid in improving the prognosis of MNCH utilisation such as classifying high-risk populations of women/children who are most likely to drop out of the CoC. This will aid in making better informed decisions in MNCH interventions targeted at improving the utilisation of the CoC.

Studies which assessed relationships between utilisation of MNCH services and neonatal outcomes such as neonatal mortality (36, 37, 94), have often used approaches that assume direct associations (36, 37, 94). Few studies have assessed interrelated (direct and indirect) pathways between MNCH services utilisation and neonatal outcomes (38, 39). Special analytical methods such as the GSEM are ideal for modelling complex interactions such as utilisation of interrelated services in MNCH and potential outcomes. GSEM allows for the simultaneous analysis of several equations, offering a versatile and comprehensive framework for examining various interactions among numerous variables. GSEM also enables the evaluation of potential mediators i.e indirect effects (81, 82). This study applies GSEM to investigate the presence of both direct and indirect associations between MNCH services utilisation and neonatal mortality. GSEM will enable a better understanding of the interconnected pathways between MNCH services utilisation and neonatal mortality for the development of more effective targeted interventions.

2.6 Overall aim

To describe the utilisation of MNCH services and their associations with neonatal outcomes using GSEM and Machine Learning models in three sub-Saharan African countries: the DRC, Kenya, and Tanzania.

2.6.1 Specific Objectives

1. To assess the trends and determinants of late ANC initiation in three sub-Saharan African countries: the DRC, Kenya, and Tanzania, from 2007 to 2016.

    Paper: Mlandu C, Matsena-Zingoni Z, Musenge E (2022) Trends and determinants of late antenatal care initiation in three East African countries, 2007–2016: A population based cross-sectional analysis. PLOS Glob Public Health 2(8): e0000534.

2. To predict the risk factors of non-utilisation of postnatal care in three sub-Saharan African countries using the Decision Tree

    Paper: Mlandu C, Matsena-Zingoni Z, Musenge E (2023). Predicting risk factors of non-utilisation of postnatal care in three sub-Saharan African countries: Application of the Decision Tree. (Under Review-*BMJ Public Health*).

3. To predict the likelihood of a mother/child dropping out from the MNCH continuum using various Machine Learning models and determining the most influential predictors in three sub-Saharan African countries.

Paper: Mlandu, C., Matsena-Zingoni, Z. & Musenge, E. Predicting the drop out from the maternal, newborn and child healthcare continuum in three East African Community countries: application of machine learning models. *BMC Med Inform Decis Mak* **23**, 191 (2023). https://doi.org/10.1186/s12911-023-02305-1.

4. To describe the mediation role of maternal, newborn and child healthcare services utilisation on neonatal mortality in three sub-Saharan African countries.

Paper: Mlandu C, Matsena-Zingoni Z, Musenge E (2023). Mediation role of maternal and newborn and child healthcare services utilisation on neonatal mortality in three sub-Saharan African countries: A generalized structural equation modeling approach. *(Under Review-Maternal Health, Neonatology and Perinatology)*

2.7 The conceptual framework describing the utilisation of MNCH services and their associations with neonatal mortality.

The conceptual framework for this study is informed by McLeroy's Socio-Ecological Model, see Figure 2 (1). McLeroy's Socio-Ecological Model considers the complex interplay of factors that operate at multiple levels to influence health services utilisation and impact health outcomes (1). The conceptual framework in this study grouped the factors at the individual, interpersonal, structural/societal and organisational or health system level. The factors at the individual level included maternal and child characteristics such as mother's age, mother's education, employment, parity, lack of autonomy, pregnancy intention, mode of delivery, exposure to mass media, child sex, child's birth weight; at the interpersonal level, relationship status; at the structural or societal level, place of residence, household head, household size, household wealth and permission to seek medical help alone; and at organisational the level, health insurance cover, travel distance to the health facility, medical costs, quality of ANC and place of delivery (1).

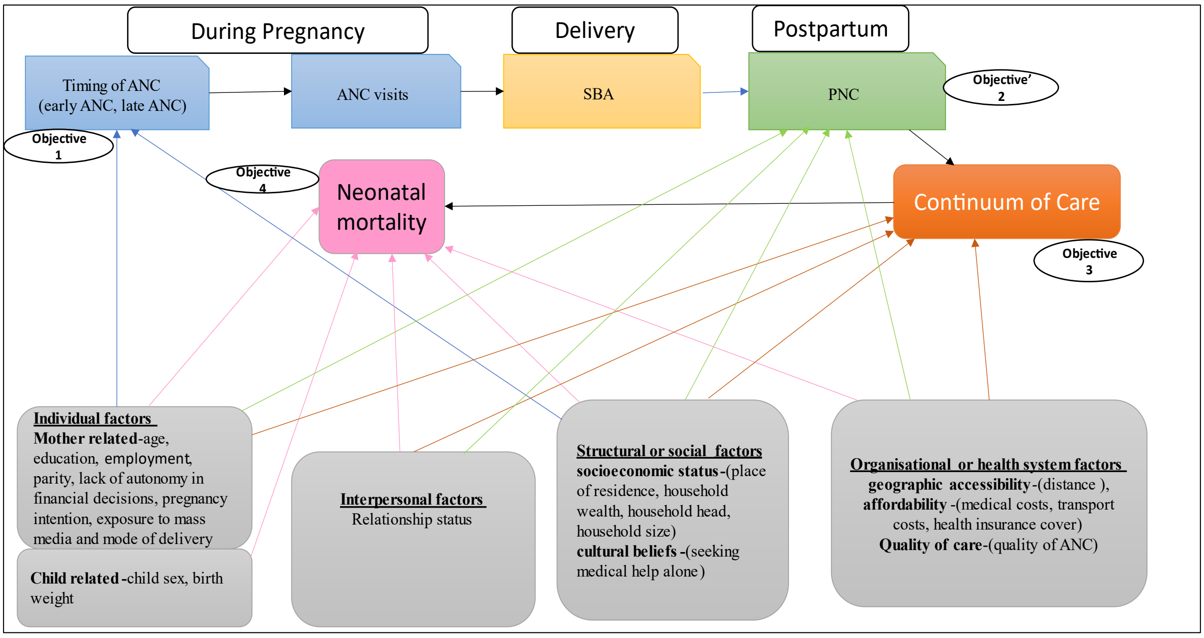

Figure 2.1: Conceptual framework describing the utilisation of MNCH services and their associations with neonatal mortality, adapted from McLeroy's Socio-Ecological Model

CHAPTER 3: METHODOLOGY

Chapter 3 describes the study setting, and design, data source, data management and the application of the statistical techniques to address the statistical and empirical research objectives of the PhD study.

3.0 Study setting

The study was conducted in three countries in SSA i.e. the DRC, Kenya and Tanzania (95, 96). These three countries were selected for the study because they rank in the top six countries in SSA with the highest number of neonatal deaths (14). The DRC is the second largest country in Africa (after Algeria), with a population of 84 million people (2018) (97). The administrative divisions of the DRC are comprised of 26 provinces (98). DRC is rich in natural resources, including minerals such as cobalt and copper, hydropower potential, significant arable land, immense biodiversity, and the world's second-largest rainforest. However, most people in DRC have not benefited from this wealth (99). A lenghty history of conflict, political upheaval and instability have contributed to the ongoing humanitarian crisis. In addition, there has been forced displacement of populations. DRC is among the five poorest countries in the world. In 2022, about 62% of Congolese, around 60 million people, lived on less than $2.15 a day (99).

Kenya's population was estimated at 46 million people in 2018 (100). Kenya is divided into 8 provinces, namely, Nairobi (the capital city), Central, Coast, Eastern, North-Eastern, Nyanza, Rift Valley (most populated) and Western provinces are subdivided into districts, and the districts are further subdivided into territories (101, 102). Over the past ten years, the nation has been through significant political and economic reforms that have boosted social development, political stability, and sustained economic growth (101). However, poverty, youth unemployment, lack of transparency and accountability, climate change continue to weaken private sector investment, and the economy's susceptibility to both internal and external shocks remain among its main development concerns (101).

Tanzania's population is estimated at 60 million (103). Tanzania is divided into 26 regions, subdivided into districts, and the districts are further subdivided into territories (104). The regions are (region capitals in parenthesis) Arusha (Arusha), Dar es Salaam (Dar es Salaam), Dodoma (Dodoma), Iringa (Iringa), Kagera (Bukoba), Kigoma (Kigoma), Kilimanjaro (Moshi), Lindi (Lindi), Manyara (Babati), Mara (Musoma), Mbeya (Mbeya), Morogoro (Morogoro), Mtwara (Mtwara), Mwanza (Mwanza), Pemba North (Wete), Pemba South (Mkoani), Pwani (Kibaha), Rukwa (Sumbawanga), Ruvuma (Songea), Shinyanga (Shinyanga), Singida (Singida), Tabora (Tabora), Tanga (Tanga), Zanzibar Central/South (Koani), Zanzibar North (Mkokotoni), and Zanzibar Urban/West (Zanzibar; Stone Town). Poverty remains largely rural, with rural areas housing four-fifths of the poor (105). Poverty is most concentrated in the western and lake zones and least concentrated in the eastern zones. Increased expansion in the agricultural sector, where the majority of the poor earn their living appears to be critical for poverty reduction (103).

3.1 Study design and data source

The study employs a repeated cross-sectional design using secondary data from the DRC (2007- 2013/14), Kenya (2008-2014), and Tanzania (2010-2015/16) DHS surveys, see Table 3.1. The DHS surveys are nationally representative cross-sectional household surveys that provide data for a wide range of monitoring and impact evaluation indicators in the areas of population, health, and nutrition (106).

Table 3.1: DHS surveys data used in the study analysis

Country	**DHS Survey year**	**Clusters**	**Total Sample**
DRC			
	2007	290	4, 559
	2013/14	492	8,941
Kenya			
	2008	386	3,512
	2014	1,573	13,776
Tanzania			
	2010	456	5,058
	2015/16	608	6,873
Combined Countries	2007-2016	3,805	42,719

3.2 Data sampling and collection

The DHS samples in the DRC, Kenya and Tanzania were stratified by urban/rural and selected in two stages. In the first stage, EAs or clusters were selected with probabilities proportional to EA size. In the second stage, households were selected per cluster with equal-probability systematic sampling. All women aged 15–49 years of the selected households or visitors who stayed in the household the night before the survey were eligible to be interviewed. The information was collected using standard questionnaires (106).

3.3 Data management and analysis

The data management and analysis is per presentation of outcomes and explanatory variables in specific objectives. More details are outlined in chapter 4,5, 6 and 7.

3.3.1 Objective 1

To assess the trends and associated determinants of late ANC initiation in three sub-Saharan African countries: the DRC, Kenya, and Tanzania, from 2007 to 2016.

The primary outcome was late ANC initiation. Women were asked how many months they were pregnant when they first received antenatal care in the pregnancy (107). The timing of ANC was categorised as early or late ANC initiation based on studies in SSA (108, 109). Late ANC initiation was coded as late if the woman's gestational age was more than four months and early if otherwise (108, 109). The explanatory variables were selected for analysis based on literature (8, 110), and they represented proxy measures for demographic and socioeconomic variables collected at the time when the woman was pregnant. These variables included place of residence, maternal age, maternal education, birth order, household size and household wealth.

Using ArcGIS, thematic maps were constructed to show the late ANC initiation prevalence rates between earlier and later surveys in the three countries. All the other analyses was conducted using STATA version 17 (StataCorp, College Station, Texas 77845 USA (111). The analysis accounted for multistage sampling. The data was svyset for clustering, survey weights and stratification prior to conducting all analyses for each round of DHS surveys. All explanatory variables were included in the analysis to allow for country and time period comparative analysis across the same explanatory variables.

The Pearson's chi-square with correction for clustering, survey weights and stratification was employed for the descriptive analysis to compare respondents' demographic and socioeconomic characteristics by late ANC initiation. The multivariate logistic regression adjusted for clustering, survey weights and stratification was used to determine the determinants of late ANC initiation. For a binary response Y_{ji} and a vector of explanatory variables X, the multivariable logistic regression model is given by:

$$\pi_{ij}(X) = \frac{exp\left(b_0 + b_1X_1 + b_2X_2 + b_pX_p\right)}{1 + exp\left(b_0 + b_1X_1 + b_2X_2 + b_pX_p\right)} \quad (1)$$

Where $\pi_{ij(X)} = P\ (Y_{jij} = 1|X)$, the probability of the j^{th} woman initiating ANC late (Y_{1ij}) in i^{th} cluster (EA) given other covariates X. Alternatively, the logit (log-odds) is linearly related to the explanatory variables, with the equation written as;

$$\text{logit}[\pi ij\ (X)] = b0 + b1\ X1ij + b2\ X2ij + \ldots + bp\ Xpij\ \ldots \quad (2)$$

for the jth woman in the ith cluster (EA).

The trends of late ANC initiation over time were assessed by comparing the earlier and later surveys and first described using proportions plotted on maps using the ArcGIS software (112) and then described using prediction scores (prediction probabilities), estimated after running the multivariate logistic regressions models accounting for clustering, survey weights and stratification, see the STATA do file in the Appendix. The differences in the prediction scores between earlier and later surveys in each country were assessed using an independent t-test.

3.3.2 Objective 2

To predict the risk factors of non-utilisation of postnatal care in three sub-Saharan African countries using the Decision Tree.

The outcome variable for this analysis was non-utilisation of PNC. The respondents were asked how long they and their newborns had a postnatal check after delivery, which could be hours, days or weeks (107). Non-utilisation of PNC was constructed into a binary variable, coded as one (1) if the mother and neonate did not utilise PNC and zero (0) if both mother and the neonate utilised PNC based on literature (113). In the DHS PNC utilisation is self-reported by the mother.

The explanatory variables were place of residence, current maternal age, wanted pregnancy, maternal education, employment in the past year, parity, exposure to mass media, relationship status, head of household, household wealth, household size, financial decision making, mode of delivery, health insurance cover, seeking medical help alone, travel distance to the health facility, medical costs, timing of ANC, ANC visits, SBA and place of delivery and quality of ANC. Quality of ANC was constructed using items on routine ANC services including measurement of weight, height and blood pressure, and collection of urine and blood samples. The items on routine ANC services were coded as binary responses (yes if the service was received and 0=no if otherwise). Women who scored ≤75% of the total score were classified as having received low quality of ANC, and high quality of ANC if otherwise.

The data was cleaned and edited using STATA/SE version 17.0 (111). The data were weighted using the "svyset" STATA command to account for clustering, survey weights and stratification. This was done to account for the effect of the hierarchical nature of DHS surveys data and to restore the representativeness of the survey. A descriptive analysis of the variables was conducted. Stepwise regression was conducted at the 5% significance to select features/explanatory variables to be considered for Machine Learning. Feature selection is an essential aspect for Machine Learning that involves choosing non-redundant and most relevant features that allows one to build optimised models of the outcome (114). The data was then imported to R software for Machine Learning to train and test the Decision Tree and Logistic regression classification models.

The data was preprocessed in R to convert it in a way that can be used to train the Machine Learning models. The selected features or explanatory viables were hot encoded or converted into factor or binary variables. The data was prepared for the Machine Learning analysis by splitting the data into training and testing data sets using an 80/20 split. The study outcome PNC non-utilisation was imbalanced. Therefore, Random Over-Sampling Examples (ROSE) was applied to the training datasets to correct the class imbalances (115). The random oversampling corrects class imbalance by replacing the training data with multiple copies of some of the minority classes (115). In this analysis, ROSE was applied to increase the number of participant who did not utilise PNC (minority class) to balance with those who utilised PNC (115).

The Decision Tree was applied to the training data to assess a combination of factors with the greatest risk of PNC non-utilisation. The Decision Tree algorithm selects the variables from the database to split the sample into progressively smaller subgroups. It first identifies the most important feature/variable variable, the root node, which divides into two branches until the next best variable is reached (78, 116), resulting in a multilevel structure that resembles a tree

(78, 116). The Decision Tree has several advantages. Firstly, the algorithm can divide consecutive data into the best predictive group, efficiently dividing populations into meaningful subgroups. Secondly, high-risk population subgroups can be identified, and interventions can be developed according to the specific needs of each subgroup, especially for vulnerable groups. Thirdly, the Decision Tree can describe associations in the data by revealing second and higher order interactions among variables. Fourthly, it gives a combination of variables with the greatest risk. The Logistic regression model was also applied to the training data to assess how well it predicts the outcome in comparison with the Decision Tree. The test data was used to validate the Decision Tree and Logistic regression models' performance, see the R code in the Appendix.

The parameters used to compare the Decision Tree and Logistic regression models were accuracy, sensitivity, specificity and area under ROC curves (117). A ROC curve is a plot of the sensitivity versus 1 − specificity of a model/test. The ROC curve represents the mean sensitivity value for a given model/test across all conceivable specificity values, or conversely. An area under the ROC curve (AUC) is an effective way to summarize the overall diagnostic accuracy of the model/test. It takes the value from 0 to 1, with 1 representing a perfectly accurate model/test and 0 representing a completely inaccurate one (118). AUC values of 0.5 indicate no discrimination, while 0.6≥AUC>0.5 indicates poor discrimination, 0.7≥AUC>0.6 indicates acceptable discrimination, 0.8≥AUC>0.7 indicates excellent discrimination, and AUC>0.9 indicates exceptional discrimination (119). The larger the AUC, the better is overall performance of the model to predict non-utilisation of PNC.

3.3.3 Objective 3

To predict the likelihood of a mother/child dropping out from the maternal, newborn and child healthcare continuum and determining the most influential predictors in three East African Community countries including the DRC, Kenya and Tanzania.

The outcome variable in this analysis was drop out from the MNCH continuum. The outcome variable was constructed using self-reported utilisation of MNCH services including ANC visits, SBA and PNC in the DHS. To measure utilisation of these services, women were asked how many times they received antenatal care during the pregnancy, who assisted with the delivery and how long they and their newborns had a postnatal check after delivery (107). The definitions for drop out in MNCH continuum were constructed using existing literature (120). Antenatal care drop out was considered if a woman had less than four ANC visits during her most recent pregnancy. Skilled birth attendance drop out was considered if a woman had four or more ANC visits but did not receive SBA (delivery was not assisted by healthcare professionals, i.e., midwives, nurses, doctors, and/or health officers). Postnatal care drop out was considered if a woman received SBA but did not attend PNC with the child within the first 6 weeks of delivery. The drop out from the MNCH continuum was coded as 1 if a woman/child either drops out of ANC, SBA, or PNC and 0 if otherwise (120).

The explanatory variables considered for analysis included the place of residence, mother's current age, mother's level of education, birth order, relationship status, exposure to mass media, medical costs, travel distance to the health facility, household size, household head and household wealth. The data was survey set using cluster, survey weights and stratification variables prior to the analysis. The univariate analysis was conducted to describe women's characteristics. Bivariate analysis was conducted to assess the women's characteristics by outcome and explanatory variables and the Pearson's chi-square was used to test the differences. Logistic regression models were fitted to identify the factors associated with drop out from the MNCH continuum. In the multivariable analysis, adjusted odds ratios (AOR) with 95% confidence intervals (CI) were used to assess the significance of the relationship between the outcome and explanatory variables.

The analysis also utilised several Machine Learning algorithms to predict the drop out from the MNCH continuum and selected the best performing Machine Learning algorithm to rank the most important features (12). Machine Learning algorithms are techniques used for big data analytics (12). These algorithms are capable of extracting information that commonly used statistical methods (logistic regression) fail to present by identifying hidden patterns and relationships in the data (12). The Machine Learning models were developed in Python 3.0 using DHS surveys data in the three countries. Features (predictors) that were found significant in the multivariable logistic regression analysis were used in the Machine Learning analysis. All the analysis in Machine Learning accounted for sample weights. The datasets were randomly assigned to the training and testing datasets using an 80/20% split. The training data consisted of the data used to develop the models and the test data or validation sets were used for evaluating the performance of the models (121).

Five classification algorithms namely the Logistic Regression, Decision Tree, Random Forest, Support Vector Machine and Artificial Neural Network were employed (80, 122). The Logistic Regression frames a binary output model with a logistic function. The logistic regression output will be a probability (0x1), which may be used to forecast the binary 0 or 1 as the output (if x0.5, output=0, else output=1) (122).

The Decision Tree algorithm seeks to classify the dataset using a model consisting of one or more trees, in the first stage. It is a tree algorithm based on transforming a complex process into a simple set of decisions, beginning at the top and proceeding downwards (122). The most basic process of training a decision tree on a dataset involves the following elements, the selection of attribute, splits in the tree, stop splitting a node and mark it terminal and the assignment of a label to each terminal node.

The Random Forest is one of the community algorithms that combines numerous decision trees. After processing the data on decision trees, an accurate estimate is tried to be produced by averaging the estimates obtained (122). Random Forest solves the overfitting problem, which is one of the most prevalent issues in traditional decision trees, by dividing both the data set and the features into multiple parts and processing them using multiple trees (122).

Support Vector Machines have emerged as highly effective Machine Learning techniques for supervised classification problems (80). Support Vector Machines are kernel-based algorithms that transform data into a high-dimensional space and build a hyperplane that maximizes the distance to the closest data point of any of the input classes (80). Although Support Vector Machines were originally developed to train binary classifiers, an extension for multiple classes is conceivable by decomposing the multiclass classification problem into several binary classification problems using one-versus-all or one-versus-one approaches (80). Artificial Neural Networks are modelled after biological neural networks found in animal or human minds. Artificial Neural Networks are composed of processing elements (PEs) with the same behavior as a biological neuron (80). Different data input, output, storage, and forwarding functions are distributed across all PEs. A number of layers (one-layer or multi-layer designs) and a number of PEs per layer comprise the layout of an Artificial Neural Network (80).

In most cases, the outcome variable is frequently unbalanced, and this affects the performance of Machine Learning models (123). The classes of the outcome variable in this study were unbalanced. Hence, random oversampling was applied to balance the distribution of classes of the outcome variable using kMeans Synthetic Minority Oversampling Technique (SMOTE)

(123). KMeans SMOTE is a oversampling method for addressing class-imbalanced data (123). It assists classification by creating minority class samples in safe and crucial areas of the input space. The approach reduces noise generation while effectively overcoming imbalances between and within classes (123).

The performance of Machine Learning models was assessed using performance metrics including Accuracy, Precision, Recall, Specificity, F1 score, as well as the AUROC (124, 125). The Accuracy metric computes the number of instances across the entire dataset in which a model made accurate predictions. However, this can be a reliable metric only if the dataset is class-balanced, which means that each class has the same number of samples (124). Precision represents the proportion of accurate positive predictions (124). Recall quantifies the proportion of actual positives for which the prediction was accurate (124). Specificity quantifies the proportion of actual negatives (126).

F1 score is a Machine Learning evaluation metric that measures a model's performance by combining the precision and recall scores of a model (124). AUROC is a performance metric for "discrimination"; it indicates a model's ability to distinguish between cases (positive examples) and non-cases (negative examples) (125).The ranking of features was performed on the Machine Learning model with better performance using the package's inbuilt functions of generating and plotting prediction scores of predictors, see the Python code in the Appendix.

3.3.4 Objective 4

To describe the mediation role of maternal, newborn and child healthcare services utilisation on neonatal mortality in three sub-Saharan African countries.

The analysis considered several endogenous and exogenous variables. Endogenous variables are variables that are determined by the model or influenced by other variables, while exogenous variables are variables that are determined outside of the model or not influenced by other variables (127). The endogenous variables included MNCH services variables; ANC attendance, coded as 1=low if ANC visits were less than four, 0=high if otherwise, SBA , coded as 1=no if the mother was not attended by a skilled worker, and 0=yes if otherwise and PNC attendance, coded as 1=no if mother and child did not have PNC within six weeks of delivery and 0=yes if otherwise. Other endogenous variables were self-reported neonatal health outcomes including LBW and neonatal mortality. In the DHS the mother was asked how much the child weighed at birth in grams from recall or using a health card if available. The mother was also asked the exact months when the child died and asked to record the time of death in days if less than 1 month, in months if less than two years or years (107). Using WHO guidelines, LBW was coded as 1=yes if the newborn weighed less than 2,500g at birth and 0=no if otherwise and neonate mortality coded as 1=yes if neonate's death was within the first month of life, and 0=no if otherwise (128, 129) .

This study utilised a Conceptual Framework informed by McLeroy's Socio-Ecological Model (SEM) (1) to select important exogenous variables based on literature, and describe potential pathways for interventions to reduce neonatal mortality, see Figure. McLeroy's Socio-Ecological Model considers the complex interplay of factors that operate at multiple levels to influence health services utilisation and impact health outcomes (1). The individual factors included maternal and child characteristics such as maternal age, parity, maternal education, exposure to mass media, wanted pregnancy, financial decision making, mode of delivery, timing of ANC, employment in past 12 months and child sex. The interpersonal factor was relationship status. Community factors comprised exposure to mass media, household wealth, household head, household size, and permission to seek medical help alone. Organisational factors comprised place of delivery, travel distance to the health facility, medical costs, health insurance cover and quality of ANC.

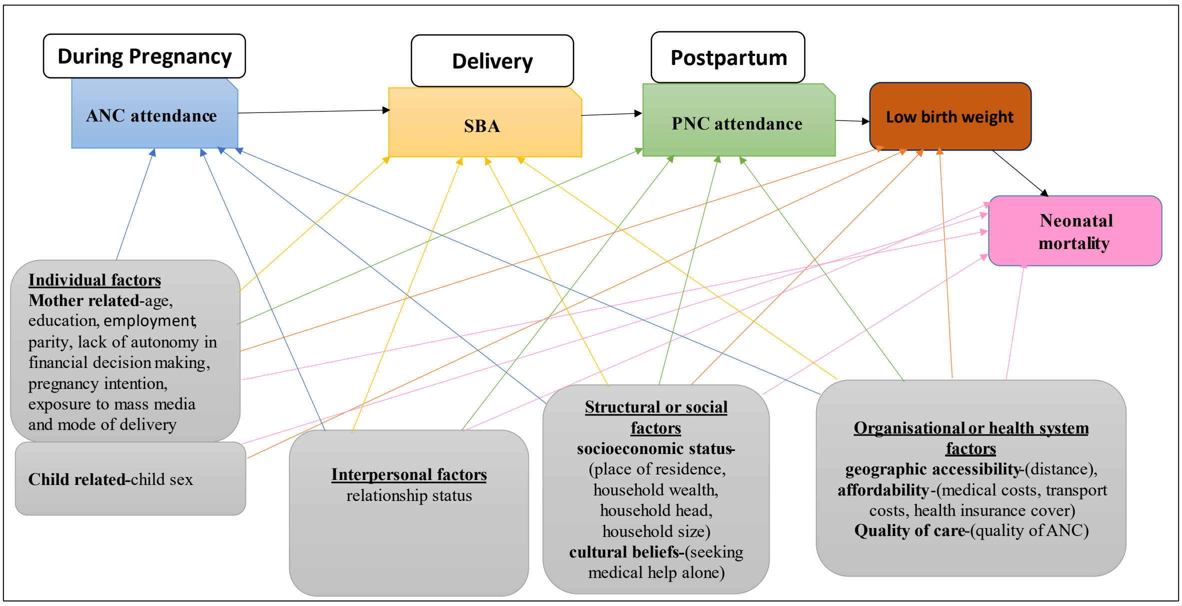

Figure 3.1: A conceptual framework for the mediation role of MNCH services utilisation on neonatal mortality [adapted from the Socio-Ecological Model (1)].

Data was analysed using Stata version 17. The data was corrected for survey design using cluster, sample weight and stratification variables before analysis. A descriptive summary of the participants' characteristics was conducted using Pearson's chi square test. To examine the MNCH services utilisation and newborn mortality pathways, several analysis steps were applied. Multivariate analysis with all covariates were performed to identify the variables that would significantly affect outcomes (endogenous variables) of interest in the adjusted models. The results were reported in terms of ORs and 95% CI. The level of significance was determined at 5%.

The Generalized Structural Equation Modelling (GSEM) with binomial logit link function (130) was used to test for the mediation role (indirect effect) of MNCH services utilisation on newborn mortality using variables that were significant in the multivariable logistic regression analysis. The advantages of utilising the GSEM compared to other statistical techniques like the conventional logistic regression or Machine Learning is that it is able to assess both direct and indirect relationships between the outcome and explanatory variable and utilises latent constructs (131). The GSEM analysis adjusted for survey weights and incorporated random effects at the cluster level. The Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) were used to assess if adding paths in the GSEM model improved the model and whether the final selected model fit with random effects at the cluster level had a better fit than the model without random effects. The BIC is more useful in selecting a correct model, while the AIC is more appropriate in selecting the best model for predicting future observations (132). The model with a lower BIC is the correct or model with the best fit (132).

The mediation analysis was performed using Stata *'gsem'* command, see the STATA code in the Appendix. A mediator is a variable that explains the relationship between the outcome and explanatory variable. Mediation can be partial or complete mediation (133, 134). Complete mediation is present when the exposure variable no longer influences the outcome variable after the mediator has been controlled (133, 134). Partial mediation occurs when the explanatory variable's influence on the outcome variable is reduced after the mediator is controlled (133, 134). Mediators can also be classified as single and sequential. A single mediator refers to when there is only one variable in the causal pathway between exposure and outcome variable as shown in Figure 3.1a. Multiple mediators are considered when more than one mediator variables operate jointly at the same stage in a causal model (135), as shown in Figure 3.1b. Thus, several indirect effects link the exposure variable to the outcome variable.

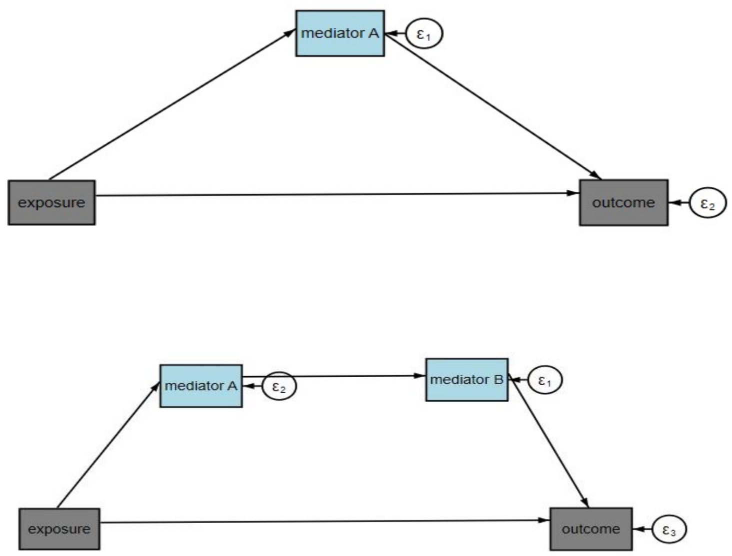

Figure 3.2a-b: a) Single Mediation b) Sequential Mediation

Stata *'nlcom'* command was used to estimate the direct, indirect, and total effects. The indirect effects were estimated using the product of the coefficients test. The "bootstrap' command was then applied to obtain bootstrap standard errors and confidence intervals. Bootstrap confidence intervals with the exclusion of "zero" showed that the indirect effects were significant.

3.5 Ethical Considerations

The ICF Institutional Review Board (IRB) ethically approved the DHS surveys. The ICF IRB approved procedures and questionnaires for standard DHS surveys. Additionally, the ICF IRB and IRB in the host country review country DHS survey protocols. ICF IRB makes sure that the DHS survey follows the United States Department of Health and Human Services regulations for the protection of human subjects, while the host country IRB makes sure that the DHS survey follows the laws and norms of the country (106). The DHS surveys were conducted in accordance with the guidelines and regulations stated in the Declaration Helsinki. Confidentiality and anonymity were ensured in the DHS surveys. The DHS surveys are conducted in an anonymous manner which does not allow any potential identification of any single household or individual in the data file. Results of interviews are kept strictly confidential. The consent statements and procedures for recording of consent and interviewers are part of the survey questionnaires. Verbal informed consent is sought by the interviewer reading a prescribed statement to the respondent and recorded in the questionnaire whether or not the respondent consented. In the case of minors, parental or guardian approval is required before a child or adolescent can participate (106).

The secondary analysis study was exempt from ethics approval because the data is publicly available. The researcher was authorised to use the DHS datasets that are publicly accessible on MEASURE DHS website https://dhsprogram.com upon request. To request for access one must be a registered user of the DHS website. Dataset requests should include contact information, a research project title, and a description of the study for the data. Once approved the registered user is granted access to the data (106).

3.6 Papers and Supplementary Results

The next section of chapters will be stand-alone published papers and work under review for each objective. These chapters have been ordered according to the 4 broad objectives of the thesis. Chapter 4 assesses the trends and determinants of late ANC initiation, while Chapter 5 describes the use of the Decision Tree in predicting the main risk factors of non-utilisation of PNC. Chapter 6 describes the use of several complex Machine Learning methods in predicting a mother/child dropping out from the MNCH continuum and determining the most influential predictors. Chapter 7 describes the mediation role of MNCH services utilisation on neonatal mortality.

CHAPTER 4: EMPIRICAL EVIDENCE for Paper/objective 1

Trends and determinants of late antenatal care initiation in three sub-Saharan African countries, 2007–2016: A population based cross-sectional analysis.

Authors: Chenai Mlandu[1], Zvifadzo Matsena-Zingoni[2,] Eustasius Musenge[2]
Affiliation: School of Public Health, University of Witwatersrand, South Africa[1,2]
(Published Global Public Health: August 15, 2022 https://doi.org/10.1371/journal.pgph.0000534)

Abstract

Background: Early antenatal care is critical for the mother and newborn's health. Antenatal care is often delayed in Sub-Saharan Africa. The study aims to examine the trends and determinants of late antenatal care initiation in the Democratic Republic of Congo, Kenya, and Tanzania from 2007-2016.

Methods: The study employed Demographic Health Surveys data of reproductive-age women seeking antenatal care in the Democratic Republic of Congo (2007-2013/14), Kenya (2008-2014), and Tanzania (2010-2015/16). Bivariate and multivariate analysis was conducted per survey, taking sampling weights into account. The determinants of late antenatal care initiation were measured using multivariate logistic regression models and the trends were assessed using prediction scores.

Results: Late antenatal care initiation declined in Tanzania (60.9%-49.8%) and Kenya (67.8%-60.5%) but increased in the Democratic Republic of Congo (56.8%-61.0%) between surveys. In the Democratic Republic of Congo, higher birth order was associated with antenatal care initiation delays from 2007-2014, whilst rural residency (AOR:1.28;95%CI:1.09-1.52), lower maternal education (AOR:1.29;95%CI:1.13-1.47) and lower-income households (AOR:1.30;95%CI:1.08-1.55) were linked to antenatal care initiation delays in 2014. In Kenya, lower maternal education and lower-income households were associated with antenatal care initiation delays from 2008-2014, whilst rural residency (AOR:1.24;95%CI:1.11-1.38) and increased birth order (AOR:1.12; 95%CI:1.01-1.28) were linked to antenatal care initiation delays in 2014. In Tanzania, higher birth order and larger households were linked to antenatal care initiation delays from 2010-2016, whilst antenatal care initiation delays were associated with lower maternal education (OR:1.51;95%CI:1.16-1.97) in 2010 and lower-income households (OR:1.45;95%CI:1.20-1.72) in 2016.

Conclusion: Except for the Democratic Republic of Congo, other countries in the region are making progress in reducing antenatal care delays. Women from various geographic, educational, parity, and economic groups exhibited varying levels of delayed antenatal care uptake. Increasing women's access to information platforms and strengthening initiatives that enhance female education, household incomes, and localise services may enhance early antenatal care utilisation.

Keywords: demographic and health survey, inequalities, late antenatal care initiation, maternal mortality and trends

4.0 Introduction

Maternal mortality reduction remains a top priority in the new Sustainable Development Goals (SDG) 3.1 (136). Maternal mortality, on the other hand, remains a global issue, with 275,288 deaths in 2015 attributed to pregnancy and related complications (137). Sub-Saharan Africa (SSA) has the highest regional maternal mortality rate in the world, accounting for 66% (201 000) of global mortality (303 000) in 2015 (137). Also, in SSA, countries like the Democratic Republic of the Congo (DRC), Kenya, and Tanzania were among the ten countries in the world with significant contributions to global mortality in 2015 (137). However, the vast majority of maternal deaths are avoidable, detectable, and treatable (3, 4).

Antenatal care (ANC) is one of the key strategies for reducing maternal mortality, both directly by detecting and treating pregnancy-related complications and indirectly by identifying women at high risk of birth issues (3, 4). To lower the risk of maternal death, ANC providers give appropriate medical and educational measures (3-5). Pregnant women should begin ANC during their first trimester, no later than 16 weeks of pregnancy, according to the World Health Organization (WHO) (5). Early ANC has also been demonstrated to reduce negative perinatal outcomes such as preterm birth, stillbirth, and low birth weight (6, 7).

Early ANC registration aids health providers in providing timely information and medicines based on the health of the expecting mother (5, 8). During the first ANC visit, health issues such as the human immunodeficiency virus (HIV) and syphilis are screened, and early discovery of these conditions enhances the health and survival of the unborn child (5, 8). Furthermore, uptake of therapies such as iron supplementation and immunisations can be life-saving for both mothers and newborn babies if initiated at the early stages of pregnancy (21, 138). Those who arrive late for ANC miss out on vital health information and interventions (21, 138, 139).

Although the World Health Organization (WHO) recommends that the first ANC visit should occur within the first 16 weeks of pregnancy (21), research undertaken in SSA countries has indicated poor early ANC uptake (8, 9, 138, 139). Low early ANC uptake has been linked to characteristics such as rural residency, lack of education and lack of information about ANC, unplanned pregnancy, low income, travel expenses, multiparity, and unemployment (8, 9, 138, 139).

Recent studies in the DRC, Kenya, and Tanzania have also found poor early ANC uptake, although the trends in late ANC uptake have not been thoroughly investigated (8). It is critical to understand how pregnant women's late care-seeking behavior has changed over time to guide strategic policy actions aimed at improving timely ANC commencement in the SDG era. Thus, this study aims to assess the trends and associated determinants of late ANC initiation in three sub-Saharan African countries: the DRC, Kenya, and Tanzania, from 2007 to 2016.

4.1 Materials and Methods

4.1.1 Study design and setting

The study employs a repeated cross-sectional design with secondary data from publicly available Demographic Health Survey (DHS) surveys in three sub-Saharan African countries, namely the DRC (2007-2014), Kenya (2008-2014/15), and Tanzania (2019-2015/16) (140, 141).

4.1.2 Inclusion and exclusion criteria

The study population consisted of DRC, Kenyan, and Tanzanian women aged 15 to 49 years, who had a live birth in the five years preceding each DHS survey between 2007 and 2016. Women who attended at least one ANC visit were included, and only information pertaining to the most recent pregnancy was used.

4.1.3 Data Source and Sampling

The DHS Program conducts a nationally representative standardized cross-sectional survey involving women aged 15–49 years (142). The DHS data is collected through a multistage sampling design, where the first stage involves the selection of EAs or clusters drawn from census files by strata, and the second stage involves the random selection of individual households within each selected EA or cluster. The probability of selecting each household differs from cluster to cluster; hence, sampling weights were considered in the analysis to account for under and oversampling and to restore the sample's representativeness (142). Two most recent rounds of DHS surveys at the time of the study in the DRC, Kenya and Tanzania were used. The latest DHS survey in each country was used as the comparative survey and the preceding survey as the baseline. A total sample size of 42,719 reproductive-age women with a live birth in the five years preceding each DHS survey was used in the analysis in the three countries. More information is provided in the supplementary material (Supplementary Table 1).

4.1.4 Study Variables

The outcome variable was late ANC initiation categorised as late if the timing of ANC was >4 months and early if otherwise (110). The explanatory variables were chosen for analysis based on research in SSA (3, 9-11) and availability in DHS surveys, and they represented proxy measures for demographic and socioeconomic explanatory variables collected at the time when the woman was pregnant. These explanatory variables included the place of residence (rural, urban), mother's current age group (young women aged 15-24, older women aged 25-49 years), mother's level of education (primary and no education, secondary and tertiary education), and household wealth status, which was re-categorised as poor, middle and rich in this study, using the household wealth index variable (poorest, poor, middle, richer and richest) in the DHS surveys data (143).

4.1.5 Data analysis

The data were managed, cleaned, coded, and analysed using STATA version 17 (StataCorp, College Station, Texas 77845 USA (144). The data was corrected for survey design using the svyset command with clustering, survey weights and stratification variables prior to conducting all analyses for each round of the DHS surveys. The Pearson’s chi-square with correction for survey design was conducted to compare respondents’ demographic and socioeconomic characteristics by late initiation of ANC. The multivariate logistic regression with correction for survey design was used to assess the determinants of late ANC initiation. The trends of late ANC initiation over time were first described using proportions plotted on maps using the ArcGIS software (145) and then described using prediction scores (prediction probabilities), estimated after running the multivariate logistic regressions with correction for survey design.

The differences in the prediction scores between the baseline and latest surveys in each country were assessed using an independent t-test. Similar, explanatory variables were considered in the multivariate logistic regressions for easy comparison between countries. Significance was set at a p-value of less than 5%.

4.2 Results

4.2.1 Descriptive analysis

Trends of late ANC initiation using proportions and prediction scores

Using proportions, the results on the map showed a declining trend in the prevalence of late ANC initiation in Kenya and Tanzania between the first and second surveys while the DRC showed an increasing trend as shown in Figures 4.1 and 4.2.

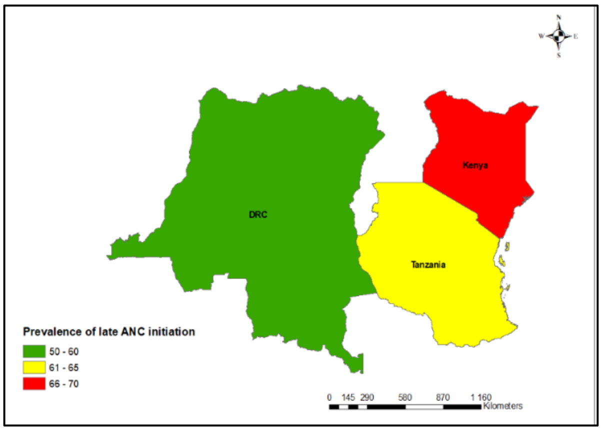

Figure 4.1.: Map showing the prevalence of late ANC initiation in the DRC, Kenya, and Tanzania in the first survey in this study: Base map source at https://spatialdata.dhsprogram.com.

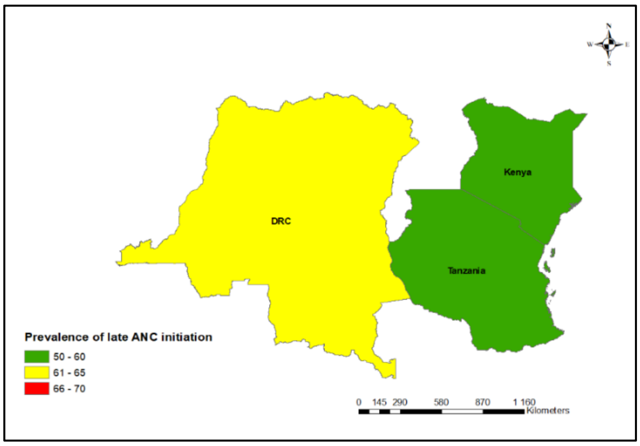

Figure 4.2: Map showing the prevalence of late ANC initiation in the DRC, Kenya, and Tanzania second survey in this study. Base map source at https://spatialdata.dhsprogram.com.

Using prediction scores, the trends revealed that Tanzania had the greatest reduction in late ANC initiation, dropping from 60.9% to 49.8% between 2010 and 2016, followed by Kenya, which dropped from 67.8% to 60.5% between 2008 and 2014. Between 2007 and 2014, the trends showed that the DRC had an increase in late ANC initiation, rising from 56.8% to 61.0% (Figure 4. 3).

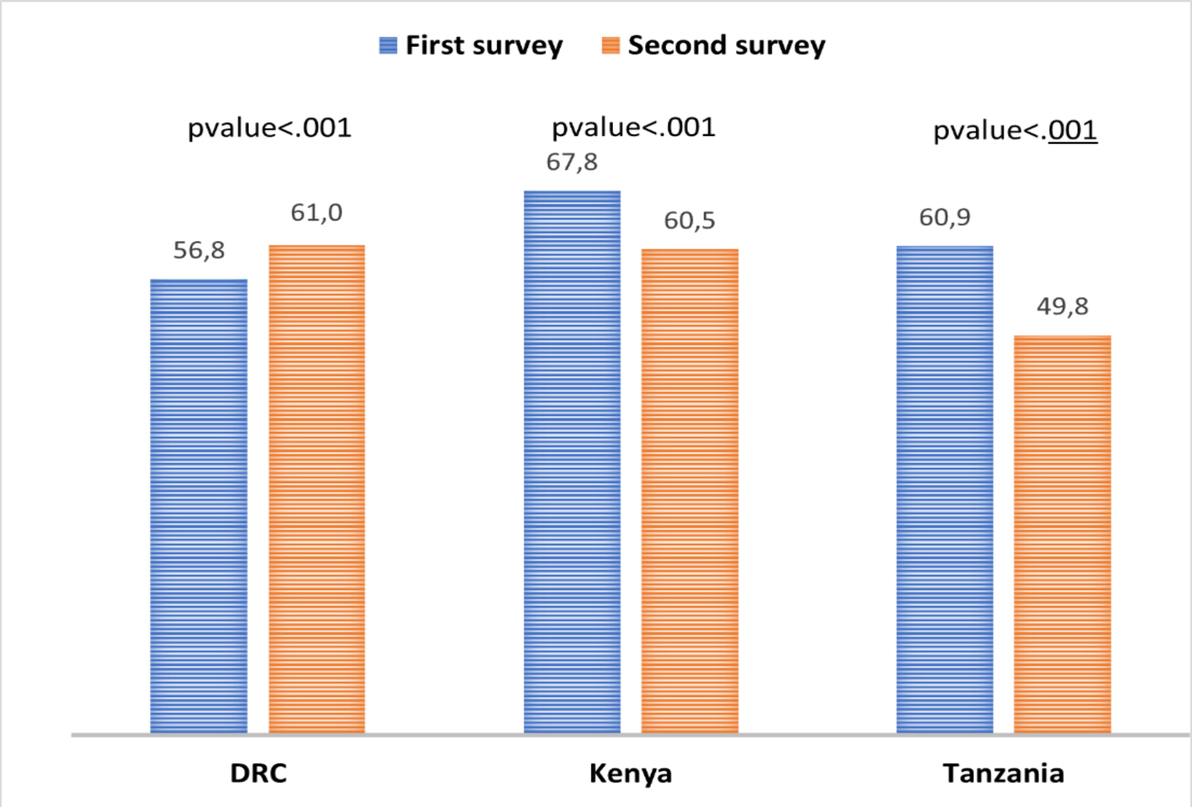

Figure 4.3: Trends of late ANC initiation in the DRC, Kenya, and Tanzania between the first and second surveys.

Late ANC initiation by background characteristics of respondents in the combined countries

Table 4.1 shows late initiation of ANC based on the demographic and socioeconomic characteristics of respondents in the combined countries' data from the first and second surveys. A total of 13,129 and 29,950 women of reproductive age who had a live birth in the five years preceding each survey were included in this study, in the first and second surveys respectively.

Table 4.1: Late ANC initiation by background characteristics in the combined countries

	Combined Countries: Late ANC initiation	
Background characteristics	First survey (n=8,242) weighted n(%)	Second survey (n=16,930) weighted n(%)
Place of residence		
Urban	2,241(56.7)	5,336(49.8)
Rural	6,001(63.5)	11,594(61.4)
	(<.001)**	(<.001)**

Age-group		
Young women (15-24 years)	2,645(59.9)	5,131(56.4)
Older women (25-49 years)	5,597(62.3)	11,799(57.6)
	(.053)	(.315)
Education		
Primary and below	6,468(63.5)	11,812(60.7)
Secondary and above	1,774(55.2)	5,118(50.6)
	(<.001)**	(<.001)**
Birth order		
One	4,288(57.7)	9,699(53.7)
Two	3,233(65.2)	5,991(61.9)
Three or more	721(70.3)	1,231(67.1)
	(<.001)**	(<.001)**
Household size		
4 or less	2,226(57.6)	5,015(51.3)
5 or more	6,016(63.0)	11,915(60.1)
	(<.001)**	(<.001)**
Household wealth		
Poor	3,242(64.2)	7,399(63.8)
Middle	1,801(64.6)	3,408(59.9)
Rich	3,016(57.1)	6,113(49.7)
	(<.001)**	(<.001)**

In the first survey, rural women initiated ANC later (63.5%) than urban women (56.7%), and a similar pattern was observed between rural women (61.4%) and urban women (49.8%) in the second survey. In the first survey, older women aged 25-49 years started ANC later (62.3%) than younger women aged 15-24 years (59.9%). In the second survey, a similar pattern was observed between older women (57.6%) and younger women (56.4%), but the differences between the two age groups were not statistically significant in either survey. In the first survey, women with primary or no education delayed their first ANC visit (63.5%) more than women with secondary or tertiary education (55.2%). In the second survey, the same pattern was observed among women with primary or no education (60.7%) and women with secondary or tertiary education (50.6%)

In the first survey, women with birth order of three or more started ANC later (70.3%) than those with birth order of two (65.2%) or one (57.7%). In the second survey, women with birth order of three or more also initiated ANC later (67.1%) than those birth order of two (61.9%) or one (53.7%). Women in households with 5 or more members started ANC later (63.0%)

than women in households with 4 or fewer members (57.6%) in the first survey. In the second survey, women in households with 5 or more members (60.1%) and women in households with 4 or fewer members (51.3%) showed a similar pattern. In the first survey, women in poor-income households delayed their first ANC visit (64.2%%) more than women in rich-income (57.1%) households. In the second survey, women in poor-income households also waited longer to start ANC (63.8%) than women in middle-income (59.9%) or rich-income (49.7%) households.

Late ANC initiation by background characteristics by country

Table 4.2 shows late initiation of ANC by demographic and socioeconomic characteristics of respondents for the various countries, i.e., DRC, Kenya, and Tanzania, in the first and second surveys. In the DRC, a total of 4,559 and 8,941 reproductive-age women with a live birth in the five years preceding each survey were included in this study in 2007 (baseline survey) and 2014 (comparative survey) respectively. In 2007, rural women started ANC later (58.7%) than urban women (54.7%), and in 2014, a similar pattern was found among rural women (64.9%) and urban women (52.5%). In 2007, however, the differences between rural and urban areas were not statistically significant.

In 2007, older women aged 25-49 years commenced ANC later (58.7%) than younger women aged 15-24 years (53.7%). In 2014, older women (61.2%) and younger women (58.9%) followed a similar pattern. However, the differences between older and younger age groups were only statistically significant in 2007. In 2007, women with primary or no education started ANC later (59.1%) than women with secondary or tertiary education (53.8%). A similar pattern was observed among women with primary or no education (65.3%) and women with secondary or tertiary education (54.2%) in 2014.

In 2007, women with birth order of three or more began ANC later (65.5%) than those with birth order of two (59.8%) or one (52.6%). In 2014, women with birth order of three or more also commenced ANC later (68.3%) than those with birth order of two (62.5%) or one (56.9%). In 2007, higher proportions of women in households with five or more members delayed their first ANC visit (57.6%) than women in households with four or fewer members (54.9%). In 2014, a similar pattern was observed for women in households with five or more members (61.3%) and women in households with four or fewer members (57.9%). In 2007, lower proportions of women in low-income families started ANC later (57.3%) than women in rich-income families (53.8%). In 2014, women in low-income families delayed ANC initiation (66.0%) more than women in middle-income (63.7%) or rich-income families (53.5%).

In Kenya, a total of 3,512 and 13,766 reproductive-age women with a live birth in the five years preceding each survey were included in this study in 2008 (baseline survey) and 2014 (comparative survey) respectively. Women in rural areas started ANC later (69.2 %) than women in urban areas (62.6%) in 2008. In 2014, a similar pattern was observed among rural women (63.8%) and urban women (51.4%). In 2008, younger women aged 15-24 years commenced ANC later (69.2%) than older women aged 25-49 years (67.0%). Younger women also started ANC later (59.1%) than older women (58.8%) in 2014. However, for both surveys, the differences in younger and older age groups were not statistically significant.

Table 4.2: Late ANC initiation by background characteristics in the DRC, Kenya and Tanzania

	DRC: Late ANC initiation		**Kenya: Late ANC initiation**		**Tanzania: Late ANC initiation**	
Background characteristics	2007(n=2,621) weighted n(%)	2014(n=5,443) weighted n(%)	2008(n=2,412) weighted n(%)	2014(n=8,067) weighted n(%)	2010(n=3,209) weighted n(%)	2016(n=3,410) weighted n(%)
Place of residence						
Urban	1,067(54.7)	1,687(52.5)	488(62.6)	2,780(51.4)	686(56.0)	870(41.7)
Rural	1,553(58.7)	3,757(64.9)	1,924(69.2)	5,287(63.8)	2,523(62.7)	2,541(53.0)
	(.233)	(<.001)**	(.018)**	(.001)**	(.013)**	(<.001)**
Age-group						
Younger women (15-24 years)	844(53.7)	1,624(58.9)	836(69.2)	2,422(59.1)	965(58.8)	1,085(48.2)
Older women (25-49 years)	1,777(58.7)	3,819(61.2)	1,576(67.0)	5,645(58.8)	2,244(62.2)	2,326(50.2)
	(.024)**	(.112)	(.319)	(.802)	(.094)	(.236)
Education						
Primary and below	1,640(59.1)	3,324(65.3)	1,810(70.5)	5,522(64.1)	3,018(62.3)	2,956(51.8)
Secondary and above	981(53.8)	2,119(54.1)	602(60.7)	2,545(50.1)	191(47.7)	454(40.3)
	(.031)**	(<.001)**	(.002)**	(<.001)**	(<.001)**	(<.001)**
Birth order						
One	1,157(52.6)	2,381(56.9)	1,413(66.8)	5,383(56.6)	1,718(56.1)	1,935(44.2)
Two	1,139(59.8)	2,399(62.5)	840(71.4)	2,323(63.5)	1,254(66.7)	1,268(58.3)
Three or more	325(65.5)	663(68.3)	159(72.4)	360(67.0)	237(76.8)	207(63.8)
	(<.001)**	(<.001)**	(.035)**	(<.001)**	(<.001)**	(<.001)**
Household size						
4 or less	609(54.9)	1,258(57.9)	844(64.5)	3,033(53.2)	773(53.5)	724(38.2)
5 or more	2,012(57.6)	4,186(61.3)	1,568(69.7)	5,033(62.9)	2,436(64.1)	2,686(53.9)
	(.367)	(.025)**	(.015)**	(<.001)**	(<.001)**	(<.001)**
Household wealth						
Poor	1,033(57.3)	2,331(66.0)	993(72.6)	3,455(66.2)	1,398(64.6)	1,613(56.8)
Middle	620(62.0)	1,161(63.7)	478(70.8)	1,583(61.9)	703(63.2)	663(50.4)
Rich	967(53.8)	1,951(53.5)	941(62.1)	3,029(51.1)	1,108(56.2)	1,134(41.6)
	(.100)	(<.001)**	(<.001)**	(<.001)**	(<.001)**	(<.001)**

In 2008, women with primary or no education started ANC later (70.5%) than women with secondary or tertiary education (60.7%). In 2014, women with primary or no education also started ANC later (64.1%) than women with secondary or tertiary education (50.1%). In 2008, mothers with birth order of three or more initiated ANC later (72.4%) than those with birth order of two (71.4%) or one (66.8%). In 2014, mothers with birth order of three or more (67.0%) also started ANC later than those with birth order of two (63.5%) or one (56.6%).

In 2008, women in families with five or more members delayed beginning ANC (69.7%) more than women in households with four or fewer members (64.5%). In 2014, a similar tendency was observed among women in families with five or more members (62.9%) and women in families with four or fewer members (53.2%). In 2008, women in poor-income households started ANC later (72.6%) than women in middle-income (70.8%) or rich-income (62.1%) households. In 2014, women in poor-income households also began ANC later (66.2%) than women in middle-income (61.9%) or rich-income households (51.1%).

In Tanzania, a total of 5,058 and 6,873 reproductive-age women were included in this study in 2010 (baseline survey) and 2016 (comparative survey) respectively. In 2010, rural women delayed their first ANC visit (62.7%) more than urban women (56.0%). In 2016, rural women also started ANC later (53.0%) than urban women (41.7%). In 2010, older women aged 25-49 years began ANC later (62,2%) than younger women aged 15-24 years (58.8%). In 2016, older women aged 25-49 years also started ANC later (50.2%) than younger women aged 15-24 years (48.2%). In both surveys, however, differences between older and younger age groups were not statistically significant.

In 2010, women with primary or no education delayed ANC initiation (62.3%) more than those with secondary or tertiary education (47.7%). In 2016, a similar pattern was observed among women with primary or no education (51.8%) and women with secondary or tertiary education

(40.3%). In 2010, mothers with birth order of three or more attended their first ANC visit later (76.8%) than those with birth order of two (66.7%) or one (56.1%). Similarly, in 2016, mothers with birth order of three or more began ANC later (63.8%) than those with birth order of two (58.3%) or one (44.2%).

In 2010, women in families with five or more members began ANC later (64.1%) than women in families with four or fewer members (53,5%). In 2016, women in families with five or more members also commenced ANC later (53.9%) than women in households with four or fewer members (38.2%). In 2010, women in poor-income households delayed their first ANC visit (64.6%) more than women in middle-income (63.2%) or rich-income households (56.2%). In 2016, women in poor-income households also started their first ANC visit (56.8%) more than women in middle-income (50.4%) or rich-income (41.6%) households.

4.2.2 Determinants of late ANC initiation in the combined countries

Table 4.3 shows the determinants of late ANC initiation in the combined countries' data in the first and second surveys, adjusted for the country of residence. Late ANC initiation was associated with the mother's education, birth order, household size, and household wealth in the first and second surveys, and the location of residence was linked to late ANC initiation in the second survey.

The results showed that the odds of late ANC initiation were significantly higher among women with primary or no education in the first survey (AOR:1.26; 95%CI:1.10-1.44) and second survey (AOR:1.33;95%CI:1.23-1.43). It was also observed that mothers with birth order of two (AOR:1.33;95%CI:1.20-1.47) and those with birth order of three or more (AOR:1.68;95%CI:1.38-2.06) had significantly greater odds of starting ANC late in the first survey. A similar observation was made in the second survey, mothers with birth order of two

(AOR:1.22;95%CI:1.14-1.31) and those with birth order of three or more (AOR:1.45;95%CI:1.26-1.66) had significantly higher odds of waiting longer to start ANC.

The results also showed that the odds of delayed ANC initiation were significantly higher among women living in households with five or more members in the first survey (AOR:1.21:95%CI:1.08-1.34) and second survey (AOR:1.31;95%CI:1.22-1.40). The study also established that women from poor-income (AOR:1.21;95%CI:1.06-1.39) and middle-income (AOR:1.29;95%CI:1.10-1.50) households had significantly greater odds of initiating ANC late in the first survey. These findings were also observed in the second survey, the odds of delaying ANC initiation were significantly higher among women from poor-income (AOR:1.34;95%CI:1.23-1.46) and middle-income (AOR:1.21;95%CI:1.10-1.33) households. The findings also revealed that women residing in rural areas had significantly greater odds of late ANC initiation (AOR:1.21;95%CI:1.12-1.31) in the second survey.

Table 4.3: Determinants of late ANC initiation in the combined countries

	Combined Countries: Late ANC initiation	
	First Survey	Second Survey
	AOR(95%CI)	AOR(95%CI)
Variables		
Country		
DRC	**Reference**	**Reference**
Kenya	1.70(1.48-1.96)**	1.04(0.96-1.12)
Tanzania	1.16(1.03-1.31)**	0.61(0.56-0.66)**
Place of residence		
Urban	**Reference**	**Reference**
Rural	0.97(0.84-1.11)	1.21(1.12-1.31)**
Age-group		
Younger women (15-24 years)	**Reference**	**Reference**
Older women (25-49 years)	1.04(0.94-1.16)	0.96(0.90-1.03)
Education		
Secondary and above	**Reference**	**Reference**
Primary and below	1.26(1.10-1.44)**	1.33(1.23-1.43)**

Birth order		
One	**Reference**	**Reference**
Two	1.33(1.20-1.47)**	1.22(1.14-1.31)**
Three or more	1.68(1.38-2.06)**	1.45(1.26-1.66)**
Household size		
4 or less	**Reference**	**Reference**
5 or more	1.21(1.08-1.34)**	1.31(1.22-1.40)**
Household wealth		
Rich	**Reference**	**Reference**
Middle	1.29(1.10-1.50)**	1.21(1.10-1.33)**
Poor	1.21(1.06-1.39)**	1.34(1.23-1.46)**

4.2.3 Determinants of late ANC initiation in the individual countries

Table 4.4 shows the determinants of late ANC initiation in the DRC, Kenya and Tanzania by survey. In the DRC, birth order was associated with late ANC initiation in 2007 and 2014, and the place of residence, mother's education, and household wealth were associated with late ANC initiation in 2014. The results showed that the odds of initiating ANC late significantly increased by 31% (95%CI:1.09-1.57) for women with birth order of two, and by 65% (95%CI:1.22-2.24) for women with birth order of three or more in 2007. Similarly, in 2014, women with birth order of two (AOR:1.16;95% CI:1.03-1.32) and women with birth order of three or more (AOR:1.51;95% CI:1.22-1.87) had significantly greater odds of late ANC initiation.

According to the study, women residing in rural areas had significantly increased odds of initiating ANC late (AOR:1.28;95%CI:1.09-1.52) in 2014. It was also established that women with primary or no education had significantly higher odds of late ANC initiation (AOR:1.29:95%CI:1.13-1.47) in 2014. The findings also revealed that women living in poor-income households had significantly increased odds of late ANC initiation (AOR:1.30; 95%CI:1.08-1.55) in 2014.

Table 4.4: Determinants of late ANC initiation in the DRC, Kenya and Tanzania

	DRC: Late ANC initiation		Kenya: Late ANC initiation		Tanzania: Late ANC initiation	
	Survey (2007)	Survey (2014)	Survey (2008)	Survey(2014)	Survey(2010)	Survey(2016)
Variables	AOR(95%CI)	AOR(95%CI)	AOR(95%CI)	AOR(95%CI)	AOR(95%CI)	AOR(95%CI)
Place of residence						
Urban	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**
Rural	1.03(0.81-1.32)	1.28(1.09-1.52)**	0.93(0.69-1.25)	1.24(1.11-1.38)**	0.96(0.78-1.19)	1.06(0.90-1.25)
Age-group						
Younger women (15-24 years)	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**
Adult women (25-49 years)	1.17(0.97-1.42)	1.00(0.87-1.15)	0.89(0.72-1.11)	0.94(0.85-1.05)	1.05(0.90-1.23)	0.94(0.83-1.07)
Education						
Secondary and above	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**
Primary and below	1.16(0.95-1.42)	1.29(1.13-1.47)**	1.33(1.04-1.70)**	1.42(1.28-1.59)**	1.51(1.16-1.97)**	1.18(0.99-1.40)
Birth order						
One	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**
Two	1.31(1.09-1.57)**	1.16(1.03-1.32)**	1.18(0.95-1.47)	1.12(1.00-1.24)**	1.43(1.23-1.66)**	1.53(1.34-1.74)**
Three or more	1.65(1.22-2.24)**	1.51(1.22-1.87)**	1.10(0.71-1.71)	1.17(0.93-1.47)	2.20(1.58-3.06)**	1.76(1.35-2.30)**
Household size						
4 or less	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**
5 or more	1.01(0.81-1.24)	1.16(1.00-1.34)	1.15(0.92-1.42)	1.26(1.14-1.39)**	1.38(1.18-1.62)**	1.65(1.44-1.90)**
Household wealth						
Rich	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**	**Reference**
Middle	1.32(1.00-1.76)	1.21(0.99-1.48)	1.38(1.01-1.88)**	1.22(1.06-1.39)**	1.14(0.92-1.41)	1.19(0.99-1.44)
Poor	1.04(0.80-1.35)	1.30(1.08-1.55)**	1.42(1.07-1.88)**	1.29(1.15-1.46)**	1.21(1.00-1.46)	1.45(1.23-1.72)**

In Kenya, the mother's education and household wealth were associated with late ANC initiation in 2008 and 2014, and the place of residence, birth order, and household size were linked to late ANC initiation in 2014. The results showed that women with primary or no education had significantly higher odds of late ANC initiation (AOR:1.33;95%CI:1.04-1.70) in 2008 and (AOR:1.42;95%CI:1.28-1.59) in 2014. The findings also showed that women in poor-income (AOR:1.42;95%CI:1.07-1.88) and middle-income (AOR:1.38; 95% CI:1.01-1.88) households had significantly higher odds of late ANC initiation in 2008. These findings were also similar in 2014, the odds of late ANC initiation were significantly higher among women in poor-income (AOR:1.29;95%CI:1.15-1.46) and middle-income (AOR:1.22;95%CI:1.06-1.39) households.

The findings also revealed that rural women had significantly higher odds of late ANC initiation (AOR:1.24;95%CI:1.11-1.38) in 2014. The study also found that women with birth order of two had significantly higher odds of delayed ANC initiation (AOR:1.12:95%CI:1.01-1.24) in 2014. Women living in households with 5 or more members also had significantly increased odds of late ANC initiation (AOR:1.38;95%CI:1.18-1.62) in 2014.

In Tanzania, late initiation of ANC was associated with birth order and household size in 2010 and 2016, and delays in ANC initiation were associated with the mother's education in 2010 and household wealth in 2016. The results showed the odds of initiating ANC late significantly increased among mothers with birth order of two (AOR:1.43;95%CI:1.23-1.66) and those with birth order of three or more (AOR:2.20;95%CI:1.58-3.16). Similarly, in 2016, the odds of starting ANC late significantly increased among mothers with birth order of two (AOR:1.53;95%CI:1.34-1.74) and those with birth order of three or more (AOR:1.76:95% CI:1.35-2.30).

The study also found that the odds of initiating ANC late significantly increased among women residing in households with five or more members (OR: 1.38; 95% CI: 1.18-1.62) in 2010 and (AOR:1.65;95%CI:1.44-1.90) in 2016. The findings also revealed that women with primary or no education had significantly increased odds of starting ANC late (AOR:1.51;95%CI:1.16-1.97) in 2010. The study also showed that women from low-income households had significantly greater odds of late ANC initiation (AOR:1.45;95%CI:1.23-1.72) in 2016.

4.3 Discussion

The present study attempts to examine the trends and determinants of late ANC initiation using multiple rounds of DHS surveys from 2007 to 2016 in the DRC, Kenya, and Tanzania. The results of this study indicated that there was a general downward trend of late ANC initiation in Kenya between 2008 and 2014 and Tanzania between 2010 and 2016, however, there was an upward trend in the DRC between 2007 and 2014. The findings indicate uneven progress in the reduction of delays in late initiation of ANC over time in Kenya and Tanzania, except for the DRC. Differences in the economic, and social environments, as well as the implementation of diverse maternal health policies, could explain this variation.

In the DRC and Kenya, rural women were more likely to delay ANC. Previous research has shown similar results (8, 146). Women in rural areas may face access issues and be less aware of health issues than women in urban areas. (146, 147). Rural women may also have additional cultural barriers prohibiting women from starting ANC early, such as seeking the spouse's approval (148). Our data also suggested increased barriers to accessing early ANC over time among rural women than urban women in the DRC and Kenya, as rural residence was a risk factor in the later surveys. The urban-rural disparities can be bridged by governments taking strong initiatives to improve access to early ANC among rural women. Among the three countries under study, Tanzania has a long standing health policy focusing on the expansion of health services in rural areas since the 1990s (149). Such initiatives which put much emphasis on rural development could improve access to maternal healthcare among rural women (149).

Consistent with previous studies, the study found that women from lower-income households were more likely to delay ANC initiation in the three countries (51, 150). High costs of care, transportation problems, and poor service provision could be affecting low-income women, resulting in ANC delays (51, 143). The provision of free maternal services and the expansion of health care facilities could improve early access to ANC for low-income women (151). Such initiatives have been adopted by the Kenyan and Tanzanian governments (149, 152, 153). However, our analysis showed that low-income women in Kenya and Tanzania continue to face more barriers to accessing early ANC than high-income women, as household wealth was a risk factor in the later surveys. Other cost-related factors like transport costs may be barriers to accessing ANC for low-income women. Thus, governments need to adopt a holistic approach in addressing barriers to accessing early ANC among low-income women (24, 25).

Our findings also revealed that women with poor educational attainment in the DRC, Kenya, and Tanzania were more likely to commence ANC late. This is corroborated by research conducted in Zambia and Ethiopia, which found that women with poor levels of education were more likely to delay ANC enrollment (147, 154). Uneducated women may have a limited understanding of ANC services and the benefits of early ANC commencement to safe pregnancy. Early ANC initiation could be improved through community interventions such as mass media campaigns and government policies promoting female education (143, 151). Our findings also indicated progress in closing differences in late access of ANC among low and high-educated women in Tanzania, as education was only a risk factor in Tanzania in the earlier survey. This improvement could be a result of increased maternal literacy rates in Tanzania from 67% in 2004 to 77% in 2016, as indicated by previous reports (93, 143).

The study also revealed that the odds of late ANC initiation increased among mothers with more previous births in the DRC, Kenya, and Tanzania. This was also reported by studies in Cameroon (155) and Uganda (156). This result could imply that high parity women are less inclined to attend ANC earlier possibly due to previous negative experiences with ANC services (157) and perceive that there is no danger of pregnancy complications because of experience with pregnancy and childbirth (158). The study also found persistent inequalities in accessing early ANC among women of low-high parity in the DRC and Tanzania and growing inequalities in Kenya, as birth order was a risk factor in the earlier and later surveys in the DRC and Tanzania, as well as in the later survey in Kenya. These findings emphasise the need to increase community interventions such as mass media campaigns, particularly for multigravida women to improve the timely use of maternal healthcare services among these women.

In Kenya and Tanzania, women with a household size of at least five members were more likely to initiate ANC late, according to the study. The findings are similar to a previous study in Ethiopia (150). Potential reasons include financial constraints due to the high expense of maintaining a large household on a small income, as well as a preoccupation with family responsibilities that result in neglect of one's health (150). Our study also showed persistent differences in access to early ANC in Tanzania between women from large and small households, as well as rising inequalities in Kenya, as household size was a risk factor in the earlier and later Tanzanian surveys, and the later survey in Kenya. Increasing community education programs through accessible media platforms and improving the proximity of health services through mobile clinics could break barriers to uptake of early ANC among women bearing responsibilities of larger households (143, 159).

4.3.1 Strengths and Limitations of the Study

The strength of this study is that the study used nationally representative data therefore the findings are generalisable and applicable to countrywide policies and interventions. The study also assessed trends and determinants of late ANC initiation in different countries in SSA, and the findings highlighted the context-specific characteristics that must be considered when establishing future policies and interventions of improving early ANC uptake. However, the study has several limitations, including recall bias due to the self-reported nature of the data and data collected for the five preceding years. Our research also depended on the completeness of data on variables of interest in all rounds of the DHS surveys; as a result, the study did not evaluate other important socioeconomic variables like employment status due to insufficient data. Due to the cross-sectional nature of the study, it was also not possible to establish temporal causality.

4.4 Conclusions

Countries like Kenya and Tanzania experienced a reduction in late initiation of ANC over time except for the DRC. Ealy uptake of ANC continues to be a challenge among rural women, women of socioeconomic status, women of high parity, and women living in large households. Public health officers should intensify efforts increasing the awareness on the benefits of early ANC uptake to all pregnant women. Governments should implement and improve policies that eliminate barriers to accessing maternal healthcare, especially among women of low economic status. These policies should entail making services accessible to women who cannot afford them and improving the provision of services at health facilities. Furthermore, government policies aimed at increasing female education would improve the early uptake of ANC.

4.5 Chapter Summary

Initiating ANC early during pregnancy is critical for the wellbeing of the mother and newborn. However, women often delay initiating ANC in SSA. Although there is evidence of individual studies on late ANC initiation, studies assessing the trends of late ANC initiation are limited to track if they have been changes over time and help inform future interventions and policies. Thus, the study aimed to assess the trends and determinants of late ANC initiation in three sub-Saharan African countries: the DRC, Kenya, and Tanzania, from 2007 to 2016. The study found declining trends in late ANC initiation in Kenya and Tanzania except for the for the DRC. The study also found differences in late ANC uptake among women from various geographic, educational, parity, and economic groups. Additional efforts in maternal healthcare which aim to increase women's access to information platforms, enhance female education, household incomes, and localise services are needed to enhance early ANC utilisation.

CHAPTER 5: EMPIRICAL EVIDENCE for paper/objective 2

The application of the Decision in predicting the main risk factors of non-utilisation of postnatal care in three sub-Saharan African countries

Authors: Chenai Mlandu[1], Zvifadzo Matsena-Zingoni[2], Eustasius Musenge[2]

Affiliation: School of Public Health, University of Witwatersrand, South Africa[1,2]

*chenai.mlandu@gmail.com

Corresponding author: Chenai Mlandu

(Under review: BMJ Public Health)

Abstract

Background: Postnatal care (PNC) is recommended for the optimum health of mothers and newborns. However, PNC uptake is poor in sub-Saharan Africa. Studies that employ data mining methods to predict the utilisation or non-utilisation of PNC are scarce. Thus, the study aims to predict the main risk factors of PNC non-utilisation in three sub-Saharan African countries using the Decision Tree.

Methods: A secondary analysis was conducted in the Democratic Republic of Congo (DRC) (2013/14), Kenya (2014) and Tanzania (2015/16) using Demographic and Health Surveys data. The Decision Tree was used to predict the non-utilisation of PNC and determine factors which confer the greatest risk. The Decision Tree models produces simple and visual graphic trees that can be easily interpreted and applied in public health. The Decision Tree's performance was compared to the Logistic Regression using Accuracy, Sensitivity, Specificity and area under the Receiver Operating Characteristics (ROC) curve.

Results: The Decision Tree models exhibited higher classification Accuracy and Sensitivity than the Logistic Regression models. The areas under the ROC curve were different for the Decision Tree and the Logistic Regression models in the DRC (61.7% vs 60.2%), Kenya (73.0% vs 74.8%) and Tanzania (65.2% vs 65.6%). Using the Decision Tree, low quality ANC, home deliveries and unemployment showed the highest probability of PNC non-utilisation

(92.0%) in the DRC. In Kenya, home deliveries, unemployment and non-access to mass media showed the highest likelihood of PNC non-utilisation (87.0%). In Tanzania, home deliveries, low quality ANC and unwanted pregnancies exhibited the highest likelihood of PNC non-utilisation (100.0%).

Conclusion: Women who had poor quality of ANC, home deliveries, unemployment, unwanted pregnancies, and no exposure to mass media were identified as high-risk populations of non-utilisation of PNC. Implementing targeted interventions among these subpopulations can ensure effective use of public health resources and improve PNC utilisation.

Keywords: Decision Tree, Machine Learning, Neonatal Mortality, Postnatal care

5.0 Introduction

The rates of child mortality are high in the initial month following birth, with more than 50% of newborn deaths occurring within the first 24 hours after delivery (160). About 2.5 million children died in the first month of life in 2017, approximately 7000 every day (2). The likelihood of survival from birth varies widely by region. Sub-Saharan Africa (SSA) had the highest Neonatal Mortality Rate (NMR) in 2017 at 27 deaths per 1000 live births, followed by Southern Asia with 26 deaths per 1,000 live births (2). Nearly half of the newborn deaths in SSA occur in six countries namely Nigeria, DRC, Ethiopia, Tanzania, Uganda and Kenya (14). The major causes of newborn deaths are preterm birth complications, intrapartum-related events, and infections (17). Most newborns deaths result from preventable and treatable causes (17).

Adequate provision of Maternal, Newborn and Child Healthcare (MNCH) is a key strategy in ending preventable causes of newborn deaths (48). PNC is a critical component in MNCH (6, 7). It constitutes of services provided to mothers and neonates right after delivery and up to

42 days of postnatal to ensure optimum health. These services include monitoring and regular assessment of the mother and child by skilled care providers and getting health information and support for beneficial practices such as exclusive breastfeeding, maternal and newborn danger indicators, and newborn care (161). Thus, mothers and newborns who delay or do not receive PNC forfeit these services and reduce their chances of survival. Newborns experience high deaths throughout the postnatal period, particularly in the first week. Reports from a previous study showed that 71.7% of newborn deaths occurred within the first week (161). The main causes of these deaths were birth asphyxia, severe infections and diarrhoea (161). Most likely, delays in or not receiving could have contributed to these causes of newborn deaths (161).

Reports from literature show that only 55.5% of the women and children utilised PNC in SSA (11, 30). Studies conducted in SSA have linked various with non-utilisation of PNC including older age, being unemployed, low maternal education, urban residency, long health facility travel distances, fewer antenatal care (ANC) visits, giving birth at home, and complications during labour (11, 162). Non-exposure to mass media, low decision-making power of women, traditional or cultural practices and religion are among the common driving factors (11, 113, 163).

Decision Trees are emerging as reliable analytical tools for decision making (116). Decision Trees allow for screening of many variables linked to the outcome and the identification of subgroups of individuals at a higher risk (116). They pose several advantages in comparison to the traditional statistical techniques such as the logistic regression. Decision Trees are a better choice to logistic regression models when the relationship between the predictors and the response is complex and non-linear, when dealing with missing values and using large and high dimensional data (116). Decision Trees produce high order important interactions, while the Logistic regression produces important interactions, without ordering.

Decision Trees have also gained more popularity among complex Machine Learning models like the Random Forest, Support Vector Machine and Artificial Neural Network because they are simple and easy to interpret (116, 164). The results of the Decision Tree model can be displayed in a simple graphical visual tree which allows users to understand and interpret the decision-making process easily (164). This is especially important in healthcare practice where transparency and explainability is crucial, making it easier for important stakeholders to trust and validate the results (164).

To date, most studies investigating factors associated with the utilisation or that non-utilisation of PNC have relied on traditional statistical techniques such as the logistic regression (11, 22, 113, 163, 165). Using the Decision Tree the study sought to predict the main risk factors of non-utilisation of PNC among reproductive-age women in the DRC, Kenya and Tanzania using the Decision Tree. The use of the Decision Tree will help to determine subpopulation groups of women with the greatest risk of not utilise PNC and assist in formulating more effective targeted interventions to improve PNC utilisation.

5.1 Methods and Materials

5.1.1 Study design and Sampling procedure

The study was conducted in three countries in SSA namely the DRC, Kenya and Tanzania. The three countries were selected for this study because they are listed in the top six countries in SSA with the highest number of neonatal deaths (14). The study utilised secondary data from cross-sectional Demographic and Health Surveys in the DRC (2013/14), Kenya (2014) and Tanzania (2015/16) which were most recent at the time of the study. The selected DHS surveys in the three countries were also comparable because the survey periods had small year gaps. The DHS samples are normally stratified by regions and urban/rural residence. The DHS surveys employed two-stage sampling design to create a balanced and representative sample of the target population. In the first stage, EAs or clusters were selected with a probability equal to size within each stratum, and in the second stage, households were selected with an equal probability from an updated household list per cluster (142). Women aged 15–49 years in selected households were eligible to be interviewed. A total of 8,940, 6,596 and 6,873 women from the DRC, Kenya and Tanzania respectively were considered for analysis.

5.1.2 Inclusion and exclusion criteria

The study comprised women reproductive aged 15-49 years who had a live birth in the past 5 years preceding DHS surveys in the DRC, Kenya and Tanzania. Only women who reported on ANC, skilled birth attendance (SBA) and PNC were included in the study. To minimise measurement and recall bias, only information related to the most recent birth was utilised.

5.1.3 Study variables

Outcome variable

The outcome variable was non-utilisation of PNC. The mother and neonate were considered to have not utilised PNC if either mother or neonate had no postnatal check within six weeks after delivery (113). It was constructed into a binary variable coded as one (1) if either mother or neonate did not utilise PNC and zero (0) if both mother and the neonate utilised PNC.

Explanatory variables

The explanatory variables were chosen based on the literature in SSA (11, 22, 113, 163), and availability in DHS surveys. The explanatory variables included demographic, socioeconomic and health service-related factors. The demographic explanatory variables comprised of place of residence (urban, rural) and maternal age (15-24 years, 25-49 years). The socioeconomic variables explanatory variables included wanted pregnancy (then, later, no more), maternal education (secondary and tertiary education, primary and no education), employment in the past year (yes, no), parity (one, two, three or more), exposure to mass media (yes, no), has a male partner (yes, no), head of household (male, female), and household wealth status, which was re-categorised as poor, middle and rich in this study, using the household wealth index variable (poorest, poor, middle, richer and richest) in the DHS surveys data (143).

Other socioeconomic explanatory variables included household size (4 or less, 5 or more), financial decision making (respondent, joint with husband/partner, husband/partner/other, has no partner), health insurance (yes, no), seeking medical help alone (no big problem, big problem), distance to health facility (no big problem, big problem) and medical costs (no big problem, big problem). Health service-related explanatory variables comprised the timing of ANC (early if gestational age is less than 4 months, late if gestational age is 4 or more months), attendance of four ANC visits (no, yes), SBA (no, yes) and place of delivery (facility, home). Quality of ANC was constructed using items on routine ANC services including measurement of weight, height and blood pressure, and collection of urine and blood samples. The items on routine ANC services were coded as binary responses (1=yes if the service was received and 0=no if otherwise). We then categorised women who scored ≤75% of the total score as having received low quality of ANC, and high quality of ANC if otherwise.

5.1.4 Data management and analysis

The data was cleaned and edited using STATA/SE version 17.0 (111). The sample's representativeness of the target population was restored by accounting for the survey design. Prior to carrying out any analysis, the data was declared as survey data by accounting for clustering, individual survey weights and stratification using the "svyset" command. A descriptive analysis of the variables was conducted using tabulations using the command "svy: tab". Stepwise regression was conducted at the 5% significance to select features/explanatory variables to be considered for Machine Learning. Feature selection is an important aspect for

Machine Learning that involves choosing non-redundant and most relevant features that allows one to build optimised models of the outcome (114). The data was then imported to R software for Machine Learning to train and test the Decision Tree and Logistic Regression classification models. The selected explanatory variables were used to build the Decision Tree and Logistic Regression classification models. The selected features or explanatory viables were hot encoded or converted into factor or binary variables.

The data was prepared for the Machine Learning analysis. by splitting the data into training and testing data sets using an 80/20 split. The study outcome PNC non-utilisation was heavily imbalanced. Machine Learning methods perform well with balanced classes of the target or outcome. Thus, Random Over-Sampling Examples (ROSE) was applied to the training datasets to correct the class imbalance by oversampling the minority class i.e replacing the training data with multiple copies of some of the minority classes (115). In this analysis, ROSE was applied to increase the number of participant who did not utilise PNC (minority class) to balance with those who utilised PNC (115). The Decision Tree was applied to the training data. The Decision Tree algorithm selects the variables from the database to split the sample into progressively smaller subgroups, resulting in a multilevel structure that resembles a tree (78, 116). When the Decision Tree algorithm identifies the most important explanatory variable, the root node, it divides into two branches until the next best variable is reached (78, 116). The Logistic Regression analysis was also conducted using the training data for comparison with the Decision Tree. The test data was used to validate the Decision Tree and Logistic Regression models' performance.

5.1.5 Comparison between Logistic Regression and Decision Tree model in Machine Learning

The parameters used to compare the Decision Tree and Logistic Regression models were accuracy, sensitivity, specificity and area under ROC curves (117). Accuracy is a metric that measures how often a Machine Learning model correctly predicts the outcome (117). Sensitivity is the ability of a Machine Learning model to classify true individuals with the outcome i.e those who did not utilise PNC (117), while specificity is the ability of a Machine Learning model to classify true individuals without the outcome (117). Sensitivity and specificity are inversely related: as sensitivity increases, specificity tends to decrease, and vice versa. Highly sensitive Machine Learning models will show more positive results for individuals with the outcome, whereas highly specific Machine Learning models will show more negative results for individuals without the outcome (117).

A ROC curve is a plot of the sensitivity versus 1 − specificity of a Machine Learning model. The ROC curve represents the mean sensitivity value for a given a Machine Learning model across all conceivable specificity values, or conversely. An area under the ROC curve (AUC) is an effective way to summarise the overall diagnostic accuracy of the Machine Learning model. It takes the value from 0 to 1, with 1 representing a perfectly accurate Machine Learning model and 0 representing a completely inaccurate one (118). AUC values of 0.5 indicate no discrimination, while 0.6≥AUC>0.5 indicates poor discrimination, 0.7≥AUC>0.6 indicates acceptable discrimination, 0.8≥AUC>0.7 indicates excellent discrimination, and AUC>0.9 indicates exceptional discrimination (119). The larger the AUC, the better the overall performance of the Machine Learning model.

5.2 Results

5.2.1 Demographic and Socioeconomic Characteristics of Study Participants

Table 5.1 presents the demographic and socioeconomic characteristics of study participants in the DRC (8,940), Kenya (6,596) and Tanzania (6,873). Most women resided in the rural areas in the DRC (64.3%), Kenya (60.1%) and Tanzania (69.7%), and were aged between 25-49 years in the DRC (69.4%), Kenya (70.8%) and Tanzania (67.3%). Majority of the women in the DRC (56.6%), Kenya (63.1%) and Tanzania (83.6%) had primary or no education. Women with three or more children were 10.8% in the DRC, 4.2% in Kenya and 4.7% in Tanzania. Women who did not want to get pregnant any more were in 6.0% in the DRC, 11.7% in Kenya and 4.7% in Tanzania. Unemployed women in the past year before the survey were 20.3% in the DRC, 28.4% in Kenya and 16.2% in Kenya.

The proportion of women without male partners was 15.3% in the DRC, 18.3% in Kenya and 19.7% in Tanzania. Majority of the households were led by males in the DRC (79.7%), Kenya (69.0%) and Tanzania (81.3%), and had of 5 or more members in the DRC (75.9), Kenya (58.3) and Tanzania (72.4). Women from poor-income households were 39.3% in the DRC, 37.5% in Kenya and 41.2% in Tanzania. Women without exposure to mass media were 54.2% in the DRC, 14.0% in Kenya and 17.1% in Tanzania. Women whose financial decisions were made by their husbands/partners/others were 35.1% in the DRC, 35.6% in Kenya and 33.3% in Tanzania.

Women with big problems of seeking medical help alone were 33.0% in the DRC, 6.2% in Kenya and 13.9% in Tanzania, those big problems with travel distance to health facilities were 39.2% in the DRC (39.2%), 24.9% in Kenya and 45.1% in Tanzania, and those with big problems of medical costs were 70.4% in the DRC, 39.1% in Kenya and 52.3% in Tanzania. Most women in the DRC (95.9%), Kenya (81.9%) and Tanzania (92.3%) had no health insurance cover.

Table 5.1: Demographic and Socioeconomic characteristics of Study Participants

	DRC N(8,940) * n(%)	Kenya N(6,596) * n(%)	Tanzania N(6,873) *n(%)
Place of residence			
Urban	3,212(35.7)	2,602(39.9)	2,083(30.3)
Rural	5,788(64.3)	3,910(60.1)	4,798(69.7)
Maternal age			
Young women (18-24 years)	2,756(30.6)	1,901(29.2)	2,249(32.7)
Older women (25-49 years)	6,244(69.4)	4,612(70.8)	4,632(67.3)
Maternal education			
Primary and below	5,086(56.6)	4,109(63.1)	5,754(83.6)
Secondary and above	3,914(43.5)	2,402(36.9)	1,127(16.4)
Employment in the past 12 months			
No	1,824(20.3)	1,851(28.4)	1,111(16.2)
Yes	7,175(79.7)	4,659(71.6)	5,770(83.8)
Birth order			
One	4,188(46.5)	4,495(69.0)	4,380(63.7)
Two	3,842(42.7)	1,744(26.8)	2,176(31.6)
Three or more	971(10.8)	273(4.2)	325(4.7)
Wanted pregnancy			
Then	6,028(67.0)	3,933(60.4)	4,558(66.3)
Late	2,425(27.0)	1,818(27.9)	1,983(28.8)
No more	546(6.0)	761(11.7)	340(4.9)
Has a male partner			
No	1,379(15.3)	1,194(18.3)	1,354(19.7)
Yes	7,621(84.7)	5,320(81.7)	5,592(80.3)
Household head			
Male	7,174(79.7)	4,491(69.0)	5,592(81.3)
Female	1,826(20.3)	2,023(31.0)	1,289(18.7)
Household wealth			
Poor	3,533(39.3)	2,443(37.5)	2,837(41.2)
Middle	1,821(20.2)	1,218(18.7)	1,316(19.1)
Rich	3,646(40.5)	2,850(43.8)	2,727(39.7)
Household size			
4 or less	2,172(24.1)	3,714(41.7)	1,898(27.6)
5 or more	6,828(75.9)	3,800(58.3)	4,983(72.4)
Exposure to media			
No	4,874(54.2)	914(14.0)	1,175(17.1)
Yes	4,126(45.8)	5,597(86.0)	5,705(82.9)
Financial decisions			
Respondent alone	687(7.7)	441(6.9)	242(3.5)
Joint with husband/partner	3,692(41.6)	2,494(38.9)	2,971(43.4)

Husband/partner/other	3,199(35.1)	2,280(35.6)	2,278(33.3)
Has no partner	1,379(15.6)	1,194(18.6)	1,354(19.8)
Seeking medical help alone			
No big problem	6,028(67.0)	6,106(93.8)	5,921(86.1)
Big problem	2,967(33.0)	403(6.2)	960(13.9)
Travel distance to health facility			
No big problem	5,471(60.8)	4,892(75.1)	3,776(54.9)
Big problem	3,524(39.2)	1,619(24.9)	3,105(45.1)
Medical costs			
No big problem	2,667(29.6)	3,963(60.9)	3,283(47.7)
Big Problem	6,330(70.4)	2,548(39.1)	3,598(52.3)
Health insurance cover			
No	8,632(95.9)	5,336(81.9)	6,351(92.3)
Yes	368(4.1)	1,176(18.1)	530(7.7)

5.2.2 Health service-related Characteristics of Study Participants

Table 5.2 presents the health service-related characteristics of study participants in the DRC (8,940), Kenya (6,596) and Tanzania (6,873). Women who initiated ANC late were 60.5% in the DRC, 58.2% in Kenya and 49.6% in Tanzania. Nearly 47% of women in the DRC (46.8%), 39% in Kenya (39.4%) and 48% in Tanzania (47.9%) attended less than four ANC visits. The proportion of women who had low quality ANC were 47.5% in the DRC, 42.0% in Kenya and 47.2% in Tanzania. A greater proportion of women in the DRC (89.3%), Kenya (84.5%) and Tanzania (66.3%) had no skilled attendance at delivery. About 12% of women in the DRC (12.0%), 30% in Kenya (30.4%) and 31% in Tanzania (31.1%) delivered at home. Women who delivered by caesarean section were 6.3% in the DRC, 9.5% in Kenya and 7.0% in Tanzania. A greater proportion of women and their children in the DRC (84.6%), Kenya (53.8%) and Tanzania (79.1%) did not utilise PNC.

Table 5.2: Health service-related characteristics of Study Participants

	DRC N(8,940) * n(%)	Kenya N(6,596) * n(%)	Tanzania N(6,873) *n(%)
Timing of ANC			
Early	3,557(39.5)	2,723(41.8)	3,470(50.4)
Late	5,443(60.5)	3,790(58.2)	3,410(49.6)
Less than four ANC visits			

No	4,765(53.2)	3,946(60.8)	3,565(52.1)
Yes	4,191(46.8)	2,545(39.2)	3,283(47.9)
ANC quality			
Low	4,278(47.5)	3,775(58.0)	3,251(47.2)
High	4,723(52.5)	2,738(42.0)	3,630(52.8)
Lacked SBA			
No	967(10.7)	1,009(15.5)	2,316(33.7)
Yes	8034(89.3)	5,502(84.5)	4,564(66.3)
Place of delivery			
Facility	7,920(12.0)	4,530(69.6)	4,742(68.9)
Home	1,080(88.0)	1,981(30.4)	2,139(31.1)
Mode of delivery			
Normal	8,424(93.7)	5,886(90.5)	6,401(93.0)
Caesarean Section	566(6.3)	620(9.5)	480(7.0)
Non-utilisation of PNC			
No	1,381(15.4)	3,007(46.2	5,442(20.9)
Yes	7,619(84.6)	3,506(53.8)	1,439(79.1)

5.2 Decision Tree Analysis

The Decision Tree analysis is displayed in Fig. 5.1-3. In the DRC, six significant explanatory variables were selected to define further branches and classify the probability of non-utilisation of PNC: quality of ANC, place of delivery, mode of delivery, maternal education, place of residence and employment status, see Fig 5.1. The Decision Tree diagram below illustrates that women who did not utilise PNC are split into eight classifications or groups by reading the results of the diagram using the top-down stopping rule, i.e. beginning with the parent node, quality of ANC. The Decision Tree algorithm chose quality of ANC as the first variable to split the root node, dividing the population into two groups (low quality of ANC and high quality of ANC), which were then segmented into eight classes. The probability of not utilising PNC for the combination of factors considered in the Decision Tree analysis ranged from 24.0% to 92.0%. Women who had high-quality ANC, lived in the urban areas and had secondary or higher education showed the lowest risk of not utilising PNC. Women who had low-quality ANC, delivered birth at home and were unemployed were classified as having the highest probability of not utilising PNC.

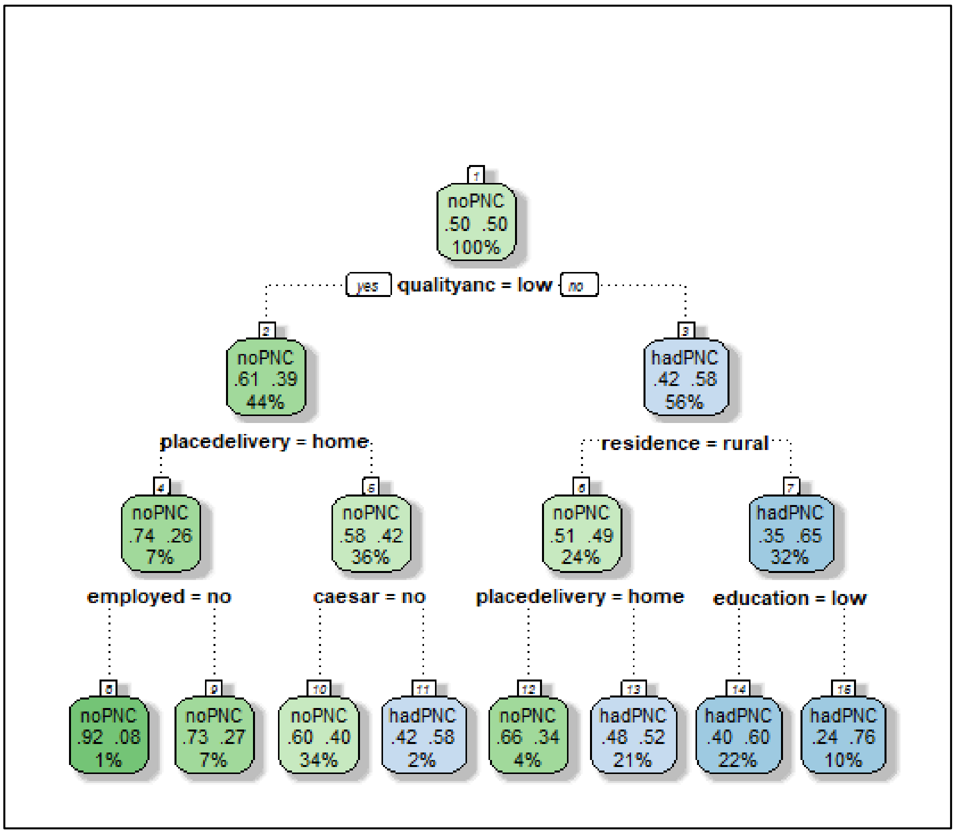

Figure 5.1: Decision Tree for predicting the risk of non-utilisation of PNC among women in the DRC; 1) root, 2) low quality ANC, 3) high quality ANC, 4) home delivery, 5) facility delivery, 6) rural residence, 7) urban residence, 8) no , 9) had employment, 10) normal delivery, 11) caesarean section, 12) home delivery, 13) facility delivery, 14) primary or no education, 15) secondary or higher education.

In Kenya, six significant explanatory variables were selected to define further branches and classify the probability of non-utilisation of PNC: place of delivery, employment status, SBA, mode of delivery, media exposure and seeking medical help alone, see Fig 5.2. The Decision Tree chose the variable place of delivery as the first variable to split the root node, dividing the population into two groups (home delivery and facility delivery), which were further separated into eight classes. The probability of not using PNC for the combination of factors considered in the Decision Tree analysis ranged from 21.0% to 87.0%. The classification with the lowest probability of non-utilisation of PNC included women who delivered in facilities, delivered by caesarean section and had SBA. Women who delivered at home, were unemployed and had no exposure to mass media had the highest likelihood of not utilising PNC.

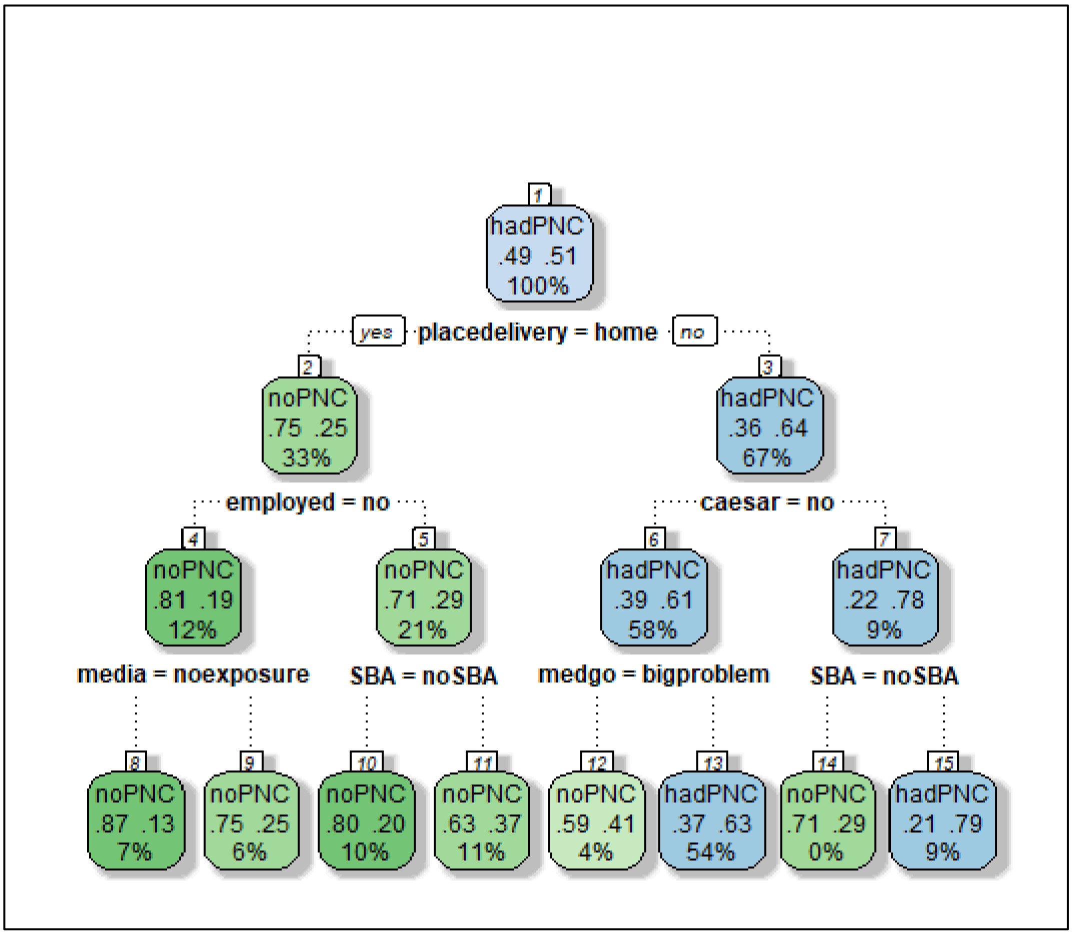

Figure 5.2.Decision Tree for predicting the risk of non-utilisation of PNC among women in Kenya; 1) root, 2) home delivery, 3) facility delivery, 4) unemployment, 5) employed 6) normal delivery, 7) caesarean section, 8) non exposure to mass media, 9) exposure to mass media, 10) no SBA, 11) had SBA, 12) big problem seeking medical help alone, 13) no big problem seeking medical help alone, 14) no SBA, 15) had SBA.

In Tanzania, five significant explanatory variables were selected to define further branches and classify the probability of non-utilisation of PNC: place of delivery, quality of ANC, seeking medical help alone, medical costs and wanted pregnancy, see Fig 5.3. The Decision Tree chose the variable place of delivery as the first variable to split the root node, dividing the population into two groups (home delivery and facility delivery), which were further separated into eight classes. The probability of not using PNC for the combination of factors considered in the Decision Tree analysis ranged from 38.0% to 100.0%. Women who delivered in facilities, received high quality ANC, and had no big problem with seeking medical help alone showed the lowest risk of not utilising PNC. Women who delivered at home, received low quality ANC and wanted pregnancy no more had the highest probability of not utilising PNC.

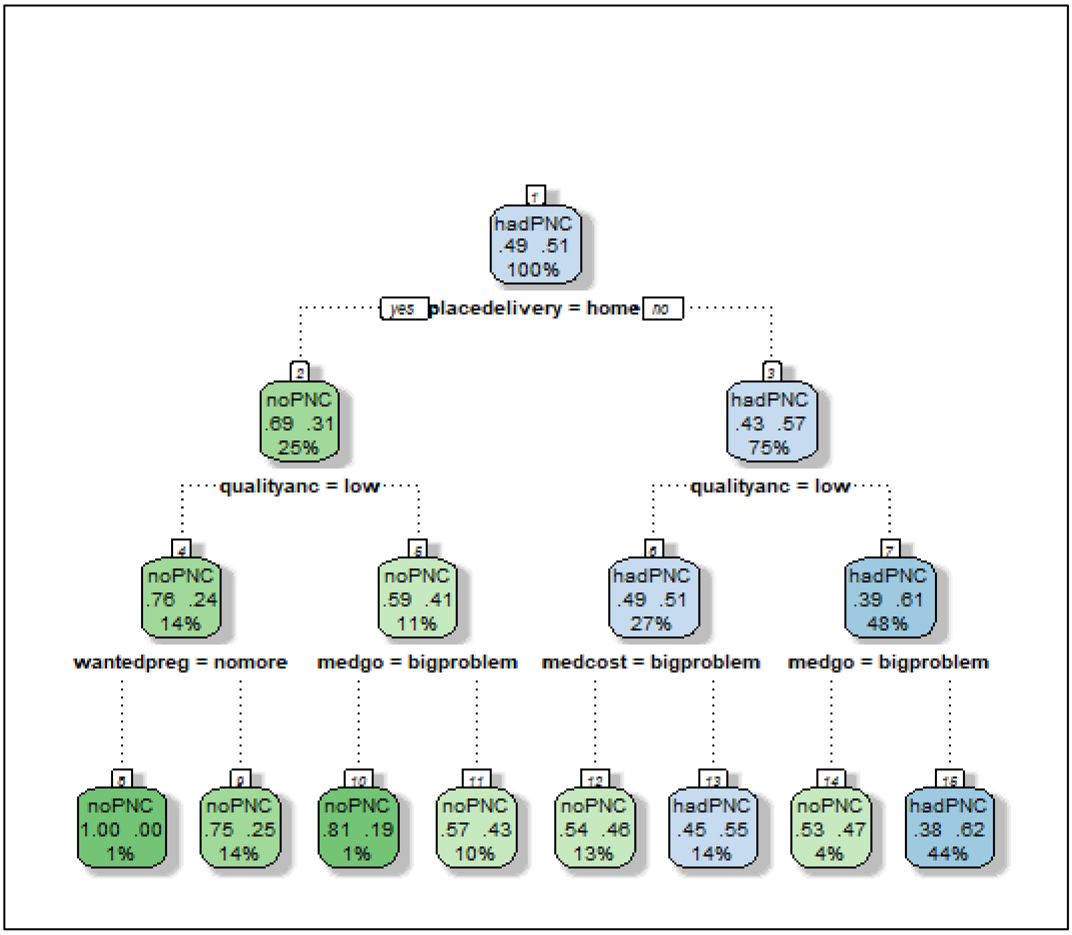

Figure 5.3: Decision Tree for predicting the risk of non-utilisation of PNC among women in Tanzania; 1) root, 2) home delivery, 3) facility delivery, 4) low quality of ANC, 5) high quality of ANC 6) low quality of ANC , 7) high quality of ANC, 8) wanted pregnancy no more, 9) wanted pregnancy then, 10) big problem with seeking medical help alone 11) no big problem with seeking medical help alone, 12) big problem with medical costs, 13) no big problem with medical costs, 14) big problem with seeking medical help alone, 15) no big problem with seeking medical help alone.

5.3 Performance of the Decision Tree and the Logistic Regression Models

The performance of Decision Tree and Logistic Regression analysis for prediction of non-utilisation of PNC was evaluated using Accuracy, Sensitivity, Specificity and AUC parameters calculated from the confusion matrices shown in Table 5.2. In the DRC, the Decision Tree model achieved a classification Accuracy of 63.1% with a Sensitivity of 64.9% and a Specificity of 53.1% while the Logistic Regression obtained a classification Accuracy of 57.8% with a Sensitivity of 58.0% and a Specificity of 57.0%. In Kenya, the Decision Tree model achieved a classification Accuracy of 67.0% with a Sensitivity of 61.0% and a Specificity of

75.6% while the Logistic Regression obtained a classification Accuracy of 66.7% with a Sensitivity of 57.2% and a Specificity of 79.6%.

In Tanzania, the Decision Tree model achieved a classification Accuracy of 61.4% with a Sensitivity of 60.3% and a Specificity of 65.8% while the Logistic Regression obtained a classification Accuracy of 57.3% with a Sensitivity of 53.6% and a Specificity of 72.0%. The AUC values for the Decision Tree and Logistic Regression models were 61.7% vs 60.2% in the DRC, 73.0% vs 74.8% in Kenya and 65.2% vs 65.6% in Tanzania. AUC values for the Decision Tree and Logistic Regression models in the DRC and Kenya showed that the models had the capacity to distinguish individuals with or without the outcome. The Decision Tree and Logistic Regression models for Tanzania showed greater capacity to distinguish individuals with or without the outcome (AUC >70%). In conclusion, the Decision Tree had better classification Accuracy and Sensitivity than the Logistic Regression in the DRC, Kenya and Tanzania but the Decision Tree had a lower Specificity than the Logistic Regression. The AUC values for the Decision Tree and Logistic Regression models showed the models had capacity to distinguish individuals with or without the outcome.

Table 5.2: Performance parameters of the Decision Tree and Logistic Regression models in the DRC, Kenya, and Tanzania

	DRC		**Kenya**		**Tanzania**	
Parameter	LR	DT	LR	DT	LR	DT
Accuracy	58.9%	61.7%	66.7%	67.0%	57.3%	61.4%
Sensitivity	59.6%	64.9%	57.2%	61.0%	53.6%	60.3%
Specificity	55.1%	53.1%	79.6%	75.6%	72.0%	65.8%
AUC	60.2%	61.7%	74.8%	73.0%	65.6%	65.2%

**DT-Decision Tree, LR-Logistic Regression

5.4 DISCUSSION

The study aimed to predict the main risk factors of non-utilisation of PNC among reproductive-age women using a Decision Tree in the three sub-Saharan African countries including the DRC, Kenya and Tanzania. The Decision Tree model had a better performance in Accuracy and Sensitivity compared to the Logistic Regression model. This means the Decision Tree had more accurate predictions of individuals who did not utilise PNC and less false positives predictions i.e it picked out more individuals who truly did not utilise PNC. Previous studies in the maternal and child health field have also found similar results (34, 35, 166).

The main advantage of Decision Tree consists of the ability to describe associations in the data by revealing high order important interactions among variables, whereas the traditional Logistic Regression shows important interactions without ordering (78). The Decision Tree results are also expressed as a tree or decision rules, which can be easily interpreted and understood (116). However, the Decision Tree like all Machine Learning algorithms does not perform well when there is imbalance of classes in the outcome variable (115). However, considering the pros and cons, Decision Tree and Logistic Regression models constitute useful and complementary analytical tools for assessing MNCH services utilisation.

In the DRC, the Decision Tree analysis showed that women who received low quality of ANC, delivered birth at home and were unemployed had the highest probability of not utilising PNC (92.0%). Quality of care refers to a range of issues, such as appropriate assessment and diagnosis, delivery of effective treatment, counseling, and user acceptance. A previous study in Ethiopia found that receiving appropriate content at ANC was associated with better health service utilisation (167). Quality of care in the DRC is generally low (168). To improve the fragile healthcare and enhance uptake of health services in DRC, health systems interventions such as the establishment of better-equipped facilities, an adequate healthcare workforce, and improved access to essential medical supplies should be established.

Women who gave birth at home in the DRC were associated with not utilising PNC. Home births in low-income countries are assisted by Traditional birth attendants (TBAs). Previous research found that most of the TBAs in low-income countries had limited formal training and access to medical treatment and equipment. Hence, women who give birth at home have a lower chance of receiving appropriate guidance from TBAs regarding the use of PNC (169). The low literacy levels and formal schooling found in most HBAs in previous research, suggests the need to train HBAs in safe birthing practices and recognition of life-threatening complications, and linking them with formal health system(169). In locations where hospitals can absorb the care for women currently delivering at home, HBAs can be given a role to direct women and their newborns to facilities to receive postnatal care (169).

Women who were not employed were also found to be associated with non-utilisation of PNC in the DRC. According to reports very few women in the DRC have access to decent jobs with 6.4% of women working in wage employment, compared to 23.9% of men (170). Women in the DRC without any means of income or gainful employment are mostly likely to face more barriers of access to maternal healthcare including transportation and medical costs, particularly those who reside in the rural areas (171). This situation in the DRC could be driven by the commercialisation of healthcare caused by fee-for-service payment systems and a lack of regulation throughout the country(172). The implementation of measures such as the government regulation of financing the health system through subsidised or free user fees could improve uptake of health services among low-income or unemployed women (172).

In Kenya, the Decision Tree analysis showed that women who delivered at home, were unemployed and had no exposure to mass media had the highest likelihood of not utilising PNC

(87.0%). Although trends show a decline in home births from 61.6% in 2009 to 52.3% in 2013, a substantial proportion of women in Kenya still give birth at home (173). According to previous studies in Kenya, home births have been linked to transportation and facility distance issues, cultural practices and these factors collectively influence women's decision to seek PNC services (11, 169, 174). Interventions such as “waiting homes” near health facilities to accommodate the expectant mothers who live far from the nearest health facilities days before delivery day can be helpful in in situations where these mothers live far (174). This will enhance their chances of getting postnatal checkups when they deliver with the assistance of skilled workers at health facilities (174).

The Decision Tree analysis in Kenya also revealed that women who were not employed had high likelihood of not utilising PNC. Long-term unemployment among women, for example, is 63.6% for women without formal education, compared to only 18.7% for men (175). Unemployment is also higher among women with an academic background beyond secondary school, at 25.8%, compared to men with the same level of education, which stands at 11.7% (175). Poverty may prevent women from seeking postnatal care because they cannot afford to pay for such services or get transport to reach health facilities (174). Although, government run health facilities in Kenya, offer free maternal health services, this may not need adequate if mothers cannot access health facilities due to distance and transport costs (174). The government needs to create strategies such as motorcycle ambulances that facilitate transportation for women of low socioeconomic who live far from health facilities and in remote areas to access health facilities and services (176).

The Decision Tree analysis in Kenya also showed that women who were not exposed to mass media were more likely not to utilise PNC. The dissemination of information via mass media has the capacity to encourage favorable behaviors and deter unfavorable health-related behaviors. The results of several prior research (177, 178) indicate that women who did not engage in weekly activities such as reading newspapers/magazines, listening to the radio, and watching television were more inclined not to utilise PNC. This could be that women with no access to mass media miss out on important information which promotes maternal and child healthcare, and thus they are less informed about PNC which decreases their chances of utilising PNC (178). Health policymakers and other non-governmental organizations should continue to invest resources in the design and implementation of education and promotion of MNCH service utilisation educational programmes via all the mass media sources to enhances MNCH uptake in SSA (178).

For Tanzania, the Decision Tree analysis showed that women who delivered at home, received low quality of ANC and had unwanted pregnancies showed the highest likelihood of not utilising PNC (100.0%). Home deliveries in Tanzania remain unacceptable high, despite the decline from 49% in 2010 to 37% in 2016 (179). A recent study in Tanzania found that poor services, distance to the nearest health facility, cost and inaccessibility of transport, sudden onset of labour and culture were amongst the reasons reported to affect home delivery (180), which all influence women's utilisation of PNC (11, 180). Interventions such as provision of financial incentives for transportation and integrating TBAs into the formal health systems could improve access to health facilities and PNC uptake.

The Decision Tree analysis also showed that women who received low quality of ANC in Tanzania were mostly likely not to utilise PNC. Tanzania is currently classified as a lower middle-income country and is endeavoring to advance to upper middle-income status (181, 182). However, the health sector faces constraints such as shortages of skilled health workers hinding its ability to enhance the quality of antenatal, delivery, and postnatal care (183). The provision of quality ANC has been linked to improved health facility seeking behavior during labour and postnatal period (183, 184). A previous study in Tanzania found that half of the health workers were not providing postnatal education during ANC visits (183) This practice could have contributed to women's low utilisation of postnatal care in this area because of not being well informed about the dangers (183, 184). Interventions such as supportive supervision, on job training to empower health workers with relevant knowledge and skills on postnatal care education could improve health facility seeking behavior during postnatal care (183, 184). Additionally, due to skilled health care workers shortages, medical attendants and community health care workers can be trained and equipped to provide postnatal care education to pregnant women with minimal supervision (183, 184).

The Decision Tree analysis also found that women who unwanted/unintended pregnancies had a high likelihood of not utilising PNC in Tanzania. These findings are similar to a previous study in Tanzania (165). Unintended pregnancies are often identified later than intended pregnancies and this could result in less time to receive ANC, affecting the subsequent utilisation of other services in the CoC (185, 186). The factors associated with lower likelihoods of PNC uptake following an unintended pregnancy vary (165, 185-187). Lower socioeconomic status, younger age, high parity have been associated with higher unintended and also reduce uptake of PNC (165, 185, 187). Targeted interventions such as improving adolescents and young women's education and improving family life planning through friendly services would reduce unintended pregnancies and improve pregnant women's decision making in seeking health services.

5.4.1 Strengths and Limitations

This study's strength is its use of nationally representative data with a sizable sample from multiple countries. Another strength is that the study employed the Decision Tree to identify high-risk subpopulations that would be targeted for maternal health interventions and optimise the use of public health resources. However, this study has limitations. The study did not evaluate other factors of PNC non-utilisation that could potentially improve the model performance such as health facility characteristics such availability of skilled care workers, quality of obstetric and postnatal care due to data unavailability. Additionally, the data on PNC was self-reported by the mother and is subject to recall bias. However, the recall bias was minimised by using information of the last birth. It is also challenging to establish a temporal relationship between the non-utilisation of PNC and the explanatory factors because the data are cross-sectional. Hence, we recommend longitudinal studies which follow up pregnant women's utilisation of services from pregnancy to postpartum period.

5.5 Conclusions

The Decision Tree is a useful and effective analytical method that can be used as a decision-making tool applied in MNCH. Using the Decision Tree, the study captured high-risk population groups of women who had the greatest risk of not utilising PNC. This one of the major advantages of the Decision Tree compared to the Logistic regression in that it captures high order interactions of variables. The interactions between variables can be represented in form of a tree which displays a combination of risk factors with the greatest risk. Using the Decision Tree, the main risk factors of not utilising PNC in the DRC were poor quality of ANC, home deliveries and unemployment. In Kenya, home deliveries, unemployment and having no access to mass media were the main risk factors of not utilising PNC. For Tanzania, home deliveries, low quality of ANC and unwanted pregnancies were the main risk factors of not utilising PNC.

To improve women's utilisation of postnatal care, government policymakers should invest resources in health financing, use of all media platforms for maternal health educational programmes, improving access to health facilities and the quality of care provided by health facilities, as well create policies which integrate TBAs into the formal health system. Additionally, healthcare providers should come up with innovative ways of addressing skilled workers shortages by training and equipping medical attendants, TBAs or community health to provide postnatal care education to pregnant women with minimal supervision and integrating maternal healthcare services with family planning services.

5.6 Chapter Summary

Most mothers and newborns did not utilise PNC after delivery in SSA. Although several studies on PNC utilisation or non-utilisation have been conducted, the use of data mining methods to predict the utilisation or non-utilisation of PNC is scarce. Thus, study aimed to predict the main risk factors of non-utilisation of PNC in the DRC, Kenya and Tanzania using a Decision Tree. The Decision Tree had better performance than the Logistic Regression. Using the Decision Tree, the study showed high-risk population groups of women who had the highest risk of not utilising PNC. This one of the major advantages of the Decision Tree compared to the Logistic regression in that it captures high order interactions of variables. The results of the Decision Tree were displayed in a visual tree which is easily interpretable. Using the Decision Tree, the main risk factors for not using PNC in DRC were low quality of ANC, home deliveries, and unemployment. In Kenya, the main risk factors for not using PNC were home deliveries, unemployment, and lack of access to mass media. In Tanzania, the main risk factors for not using PNC were home deliveries, low quality of ANC, and unwanted pregnancies.

To enhance the uptake of postnatal care, government policymakers should put more resources in health financing, use of all media platforms for maternal health educational programmes, improving access to health facilities and the quality of care provided by health facilities, and come with policies which integrate TBAs into the formal health system. Furthermore, healthcare providers should come up with innovative ways of addressing shortages of skilled workers by training and equipping medical attendants, TBAs or community health to provide postnatal care education to pregnant women with minimal supervision and integrating maternal healthcare services with family planning services.

CHAPTER 6-EMPIRICAL EVIDENCE for objective 3

Predicting the drop out from the maternal, newborn and child healthcare continuum in three sub-Saharan African countries: Application of machine Learning models.

Authors: Chenai Mlandu1, Zvifadzo Matsena-Zingoni2, Eustasius Musenge2

Affiliation: School of Public Health, University of Witwatersrand, Johannesburg, South Africa1,2

(Published *BMC Med Inform Decis Mak* **23**, 191 (2023). https://doi.org/10.1186/s12911-023-02305-1)

Abstract

Background: For optimal health, the maternal, newborn, and child healthcare (MNCH) continuum necessitates that the mother/child receive the full package of antenatal, intrapartum, and postnatal care. In Sub-Saharan Africa, dropping out from the MNCH continuum remains a challenge. Using Machine Learning, the study sought to forecast the MNCH continuum drop out and determine important predictors in three sub-Saharan African countries.

Methods: The study utilised Demographic Health Surveys data from the Democratic Republic of Congo (DRC) (2013/14), Kenya (2014) and Tanzania (2015/16). STATA 17 was used to perform the multivariate logistic regression. Python 3.0 was used to build five machine Learning classification models namely the Logistic Regression, Random Forest, Decision Tree, Support Vector Machine and Artificial Neural Network. Performance of the models was assessed using Accuracy, Precision, Recall, Specificity, F1 score and area under the Receiver Operating Characteristics (AUROC).

Result: The prevalence of the drop out from the MNCH continuum was 91.0% in the DRC, 72.4% in Kenya and 93.6% in Tanzania. Living in the rural areas significantly increased the odds of dropping out from the MNCH continuum in the DRC (AOR:1.76;95%CI:1.30-2.38), Kenya (AOR:1.23;95%CI:1.03-1.47) and Tanzania (AOR:1.41;95%CI:1.01-1.97). Lower maternal education also conferred a significant increase in the DRC (AOR:2.16;95%CI:1.67-2.79), Kenya (AOR:1.56;95%CI:1.30-1.84) and Tanzania (AOR:1.70;95%CI:1.24-2.34). Non exposure to mass media also conferred a significant positive influence in the DRC (AOR:1.49;95%CI:1.15-1.95), Kenya (AOR:1.46;95%CI:1.19-1.80) and Tanzania (AOR:1.65;95%CI:1.13-2.40). The Random Forest exhibited superior predictive accuracy (Accuracy=75.7%, Precision=79.7%, Recall=92.1%, Specificity=51.6%, F1 score=85.1%, AUROC=70%). The top four predictors with the greatest influence were household wealth, place of residence, maternal education and exposure to mass media.

Conclusions: The MNCH continuum dropout rate is very high in the DRC, Kenya and Tanzania. Maternal education, place of residence, and mass media exposure were common contributing factors to the drop out from MNCH continuum. The Random Forest had the highest predictive accuracy. Household wealth, place of residence, maternal education and exposure to mass media were ranked among the top four features with significant influence. The findings of this study can be used to support evidence-based decisions in MNCH interventions and to develop web-based services to improve continuity of care retention.

Keywords: Continuum of Care, Maternal Newborn and Child Healthcare, Machine Learning

6.0 Introduction

In the era of the Sustainable Development Goals (SDGs), reducing the global burden of preventable maternal, newborn, and child mortality and morbidity is a top priority (188, 189). Despite gains in maternal and child health during the Millennium Development Goals (MDGs) era, over 2.7 million mothers and newborn babies died in 2017 (190, 191). Sub-Saharan Africa (SSA) alone accounted for 66% (196 000) of maternal deaths, 39% of neonatal deaths

(999 000) and infant deaths in 2017 (190, 191). Most maternal and neonatal deaths occur due to avoidable complications and illnesses during pregnancy and childbirth (192, 193).

The concept of the Continuum of Care (CoC) has been brought to light to enhance Maternal, newborn and child healthcare (MNCH) through integrated service delivery (7, 188). An effective CoC connects critical MNCH packages throughout the pregnancy, delivery, and postpartum stages. Completing the MNCH CoC helps achieve the SDG 3 goals by reducing severe maternal and neonatal morbidity rates, mortality rates, and long-term physical and psychological complications (120). For instance, antenatal care (ANC) visits can identify and treat problems during pregnancy and increase the mother's chances of receiving appropriate care at birth (21, 194, 195). Skilled care during labour and delivery ensures safe and healthy delivery and reduces the risk of death for both the mother and baby (65). Postnatal care (PNC) is also recommended at birth and extends up to six weeks to avoid postpartum haemorrhage and other causes of maternal and neonatal mortality (196). A lack of appropriate care at any stage of the CoC will lead to poor MNCH outcomes.

Studies conducted in SSA including Kenya and Tanzania showed that only 10-34% of the women received complete packages of maternal health care services (197, 198). Factors such as place of residence, maternal education, maternal age, parity, household wealth, media exposure, travel distance and mode of transport have been factors found to be associated with the drop out from the CoC (31, 64, 70, 120, 197-199). Knowledge from these studies can be applied to develop Machine Learning models that can predict the likelihood of a mother/child not completing the continuum of MNCH and identify predictors with significant influence on discontinuity of care. This information will aid in the development of targeted interventions to improve MNCH retention.

Previously conventional analytical approaches such as logistic regression models have commonly been used to analyse maternal healthcare utilisation data (34). However, Machine Learning methods such as the Decision Tree and Random Forest have the potential to outperform conventional statistical methods because of various attributes including the ability to handle large and non-linear complex data, non-reliance on prior assumptions and multiple interactions between predictors (34, 200). Thus, Machine Learning methods have been reported to produce better fitting models than conventional logistic regression models (34, 121, 200). Although Machine Learning methods have better prediction performance than conventional statistical methods when applied to large datasets, studies on the application of Machine Learning methods in SSA countries remain uncommon (34, 200). Thus, this study aimed to predict the likelihood of a mother/child dropping out from the MNCH continuum by applying reliable Machine Learning predictive models and determining the most influential predictors in three sub-Saharan African countries including the DRC, Kenya and Tanzania.

6.1 Methods

6.1.1 Study Design and Settings

This study utilised secondary data from the last rounds of Demographic Health Survey (DHS) surveys in the DRC (2013/14), Kenya (2014) and 2015/16 Tanzania (2015/16) DHS surveys. The DRC, Kenya and Tanzania are neighbouring countries in the SSA (140, 141).

6.1.2 Study Population

The study comprised reproductive women aged 15-49 years who delivered their children in the past 5 years preceding DHS surveys in the DRC, Kenya, and Tanzania. Only women who attended at least one ANC visit and gave responses on skilled birth attendance (SBA) and PNC were included in the study, and only information concerning the last birth in the last 5 years was used.

6.1.3 Data Source and Sampling

The DHS is a nationally representative survey of household samples that provides comprehensive information on the population and health including MNCH. The DHS utilises multistage sampling, where the first stage involves the selection of EAs or clusters drawn from census files by strata. The second stage involves the random selection of individual households within each selected EA or cluster and the probability of selection of each household differs from cluster to cluster. The analysis used total sample sizes of 8,545, 6,432 and 6,664 reproductive-age women and children born in the past five years preceding each country's DHS survey in the DRC, Kenya and Tanzania respectively. The combined total sample for the three countries was 21,641.

6.1.4 Measurement of Variables

The outcome variable in this study was drop out from the MNCH continuum. Antenatal care drop out was considered if a woman had less than four ANC visits during her most recent pregnancy. Skilled birth attendance drop out was considered if a woman had four or more ANC visits but did not receive SBA (delivery was not assisted by healthcare professionals, i.e., midwives, nurses, doctors, and/or health officers). Postnatal care drop out was considered if a woman received SBA but did not attend PNC with the child within the first 6 weeks of delivery. The drop out from the MNCH continuum was coded as 1 if a woman/child either drops out of ANC, SBA, or PNC and 0 if otherwise (120).

The explanatory variables considered for analysis included demographic and socioeconomic variables. The demographic and socioeconomic variables included the place of residence (rural/urban), mother's current age group (15-24 years/40-44 years/25-49 years), mother's level of education (no education/primary, secondary and tertiary), birth order (1,2,3 or more), relationship status (no current partner/ has a current partner), exposure to mass media (no, yes), access to money for medication (no big problem, big problem), travel distance for medication (no big problem, big problem), household size ($<4/ \geq 5$), household head (male/ female) and household wealth (poor/middle/rich), and household wealth status, which was grouped into tertiles (poor, middle and rich) in this study, using the household wealth index variable (poorest, poor, middle, richer and richest) in the DHS surveys data (143).

6.1.5 Statistical analysis

Using the STATA package, data cleaning was performed to prepare the dataset for analysis. All analyses accounted for survey design using the svyset command with for clustering, survey weights and stratification variables. The univariate analysis was conducted to describe women's characteristics. Bivariate analysis was conducted to assess the women's characteristics by outcome and explanatory variables and the Pearson chi-square was used to test the differences. The multivariate logistic regression models were fitted to identify the factors associated with the drop out from the MNCH continuum. In the multivariable analysis, adjusted odds ratios (AOR) with 95% confidence intervals (CI) were used to assess the significance of the relationship between the outcome variables and the explanatory variables.

Machine Learning predictive models were built and trained in Python 3.0 using combined DHS surveys data for the three countries. Predictors that were found significant in the multivariable logistic regression analysis were used in the Machine Learning analysis. The Machine Learning analysis utilised the classification method. The datasets were randomly assigned to the training and testing datasets using an 80/20% split. The training data consisted of the data used to develop the models and the test data or validation sets were used for evaluating the performance of the models (121).

Five classification algorithms namely the Logistic Regression, Random Forest, Decision Tree, Support Vector Machine and Artificial Neural Network were employed. In this study, the outcome variable (drop out from the MNCH continuum) were disproportionate. Most Machine Learning algorithms work best when the number of samples in each class is about equal because most algorithms are designed to maximize accuracy and reduce errors. Thus, random oversampling was conducted to balance the distribution of classes of the drop out from the MNCH continuum. Random oversampling involves supplementing the training data with multiple copies of some of the minority classes. In this study, K-means Synthetic Minority Oversampling Technique (SMOTE) was employed to correct the class imbalance (201).

The performance of the Machine Learning predictive models was assessed using model prediction accuracies including Accuracy, Precision, Recall, Specificity, F1 score, as well as the AUROC. The ranking of features was conducted on the Machine Learning model with better performance using the feature importance permutations technique. This technique breaks the relationship between the feature and the target. The drop in the model score shows how much the model depends on the feature (79).

6.2 Results

6.2.1 Characteristics of the women

Table 6.1 shows the characteristics of women in the three countries under study. Most women in the DRC (64.2%), Kenya (60.2%) and Tanzania (69.9%) lived in rural areas. Over two-thirds of the women in the DRC (69.8%), Kenya (70.8%) and Tanzania (67.4%) were aged between 25 and 49 years. Most of the women in the DRC (85.0%), Kenya (81.7%) and Tanzania (80.7%) had a current partner. A greater proportion of the women had primary or no education in the DRC (56.5%), Kenya (63.1%) and Tanzania (83.6%). Women with birth order of three or more were about 11% in the DRC (10.6%), 4% in Kenya (4.2%) and 5% in Tanzania (4.7%). A greater proportion of the women in Kenya (79.6%) and Tanzania (67.0%) were exposed to mass media, whilst a greater proportion of the women in the DRC (87.1%) were not exposed to mass media.

A greater proportion of women in the DRC (70.2%) and Tanzania (52.4%) experienced big problems with access to money for medication, on the other hand a greater proportion of women in Kenya (60.9%) did not experience problems with access to money for medication. The majority of the women in the DRC (60.7%), Kenya (75.2%) and Tanzania (54.9%) experienced big problems with travel distance for medication. Most households in the DRC (80.0%), Kenya (68.9%) and Tanzania (81.4%) were led by males. The majority of the women in the DRC (76.6%), Kenya (58.8%) and Tanzania (73.0%%) had households with 5 or more members. Over one-third of the women in the DRC (39.3%), Kenya (37.5%) and Tanzania (41.6%) were from poor-income households.

Table 6.1: Characteristics of women in countries under study

Variables	DRC N(8,545) * n(%)	Kenya N(6,432) * n(%)	Tanzania N(6,664) *n(%)
Place of residence			
Urban	3,083(35.8)	2,524(39.8)	2,004(30.1)
Rural	5,539(64.2)	3,820(60.2)	4,663(69.9)
Maternal age-group			
Young women (15-24 years)	2,607(30.2)	1,849(29.2)	2,175(32.6)
Older women (25-49 years)	6,015(69.8)	4,495(70.8)	4,492(67.4)
Relationship status			
Has no current partner	1,295(15.0)	1,161(18.3)	1,287(19.3)
Has current partner	7,327(85.0)	5,183(81.7)	5,380(80.7)
Maternal education			
Primary and below	4,870(56.5)	4,005(63.1)	5,573(83.6)
Secondary and above	3,752(43.5)	2,339(36.9)	1,093(16.4)
Birth order			
One	3,998(46.4)	4,392(69.2)	4,232(63.5)
Two	3,711(43.0)	1,690(26.6)	2,120(31.8)
Three or more	913(10.6)	262(4.2)	315(4.7)
Exposure to mass media			
No	7,497(87.1)	1,296(20.4)	2,201(33.0)
Yes	1,115(12.9)	5,048(79.6)	4,465(67.0)
Access to money for medication			
No big problem	2,568(29.8)	3,867(60.9)	3,176(47.6)
Big problem	6,052(70.2)	2,474(39.1)	3,491(52.4)
Travel distance for medication			
No big problem	5,228(60.7)	4,769(75.2)	3,658(54.9)
Big problem	3,391(39.3)	1,573(24.8)	3,009(45.1)

Household head			
Male	6,895(80.0)	4,373(68.9)	5,429(81.4)
Female	1,727(20.0)	1,971(31.1)	1,238(18.6)
Household size			
4 or less	2,021(23.4)	2,617(41.3)	1,803(27.0)
5 or more	6,601(76.6)	3,727(58.8)	4.864(73.0)
Household wealth			
Poor	3,391(39.3)	2,379(37.5)	2,772(41.6)
Middle	1,724(20.0)	1,180(18.6)	1,258(18.9)
Rich	3,507(40.7)	2,784(43.4)	2,636(39.5)

* n(%) weighted counts and proportions of the variable

6.2.2 Patterns of the drop out from the MNCH continuum

Figure 6.1 compares the drop out from the MNCH continuum across the three countries under study. The largest gap and contributor to the drop out from the CoC occurred during the postpartum period in the DRC (83.7%), Kenya (42.4%) and Tanzania (89.4%). The overall drop out from the MNCH continuum was very high across the three countries, with proportions of 91.0% in the DRC, 72.4% in Kenya and 95.5% in Tanzania.

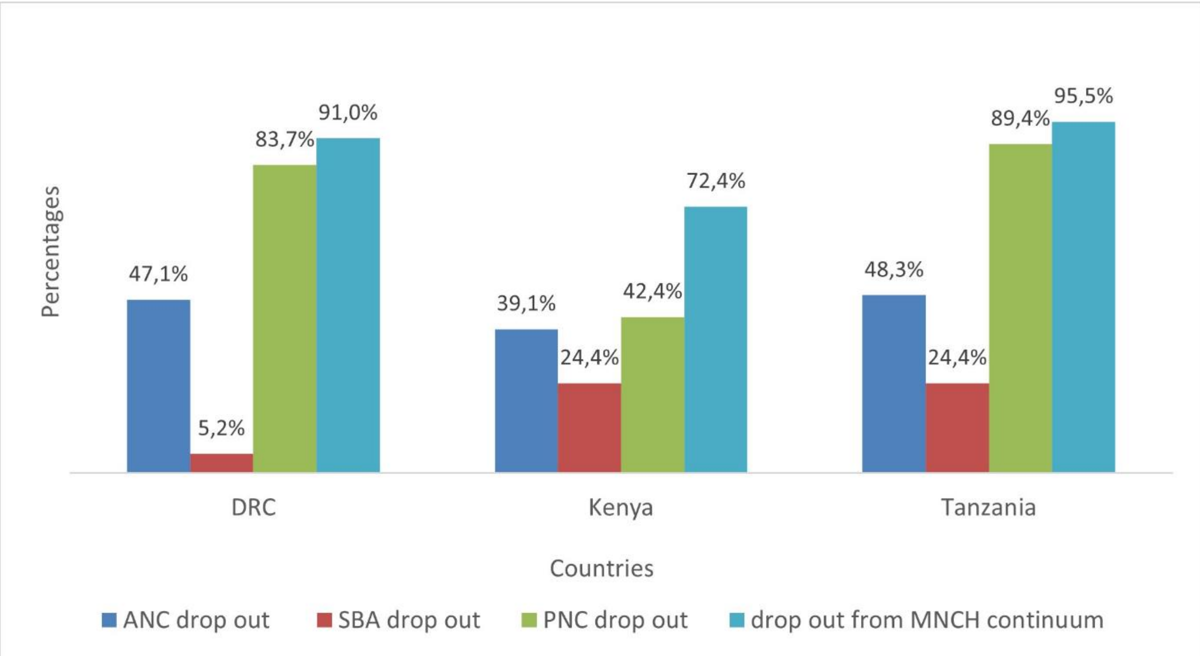

Figure 6.1: Patterns of the drop out from the MNCH continuum among reproductive-age women in the last 5 years preceding each country's DHS survey.

Table 6.2 shows results of the factors associated with the drop out from the MNCH continuum in the DRC, Kenya and Tanzania. The place of residence, mother's education and exposure to mass media were common factors significantly associated with the drop out from the MNCH continuum across the three countries. The study showed that living in rural areas, having a lower education, and having no exposure to mass media was positively associated with the drop out from the MNCH continuum.

Further, the influence of factors such as maternal age group, birth order, access to money for medication, travel distance for medication, household wealth and household size varied by country. The findings showed that the odds of dropping out from the MNCH continuum was significantly lower among older women aged 25-49 years in the DRC (AOR:0.74;95%CI:0.57-0.95) and Kenya (AOR:0.81;95%CI:0.66-0.98). An increase in birth order was also significantly associated with dropping out from the MNCH continuum in the DRC and Kenya.

The study also found that women who experienced big problems with access to money for medication (AOR:1.23;95%CI:1.03-1.46) and travel distance for medication (AOR:1.25;95%CI:1.02-1.52) had significantly increased odds of dropping out from the MNCH continuum in Kenya. Women belonging to large households also had significantly higher odds of dropping out from the MNCH continuum in Kenya (AOR:1.45;95%CI:1.22-1.72). It was also observed that women from poor-income and middle-income households had significantly increased odds of dropping out from the MNCH continuum in Kenya and Tanzania.

Table 6.2: Factors associated with the drop out from the MNCH continuum in the DRC, Kenya and Tanzania

	DRC			Kenya			Tanzania		
	*n(%)	Univariate OR(95%CI)	Adjusted OR(95%CI)	*n(%)	Univariate OR(95%CI)	Adjusted OR(95%CI)	*n(%)	Univariate OR(95%CI)	Adjusted OR(95%CI)
Place of residence									
Urban	2,633(85.4))	Reference	Reference	1,599(61.8)	Reference	Reference	1,836(91.6)	Reference	Reference
Rural	5,213(94.1)	2.73(2.20-3.39)**	1.76(1.30-2.38)**	3,033(79.4)	2.39(2.04-2.79)**	1.23(1.03-1.47)**	4,532(97.2)	3.17(2.39-4.19)**	1.41(1.01-1.97)**
Maternal age-group									
Young women (15-24 years)	2,396(91.9)	Reference	Reference	1,378(74.6)	Reference	Reference	2,084(95.8)	Reference	Reference
Older women (25-49 years)	5,450(90.6)	0.85(0.67-1.07)	0.74(0.57-0.95)**	3,212(71.5)	0.86(0.721-1.01)	0.81(0.66-0.98)**	4,282(95.4)	0.90(0.67-1.22)	0.89(0.66-1.20)
Relationship status									
Has current partner	6,672(91.2)	Reference	Reference	3,761(72.6)	Reference	Reference	5,218(95.3)	Reference	Reference
No current partner	1,172(90.5)	0.94(0.70-1.25)	1.05(0.76-1.45)	831(71.6)	0.95(0.78-1.16)	0.98(0.78-1.23)	1,240(96.4)	1.30(0.89-1.90)	1.50(0.97-2.32)
Maternal education									
Secondary and above	4,620(86.0)	Reference	Reference	1,400(79.7)	Reference	Reference	987(90.3)	Reference	Reference
Primary and below	3,225(95.0)	3.03(2.44-3.75)**	2.16(1.67-2.79)**	3,192(59.9)	2.63(2.25-3.09)**	1.56(1.30-1.84)**	5,381(96.6)	3.02(2.24-4.10)**	1.70(1.24-2.34)**
Birth order									
One	3,550(88.0)	Reference	Reference	3,002(68.3)	Reference	Reference	4,006(94.7)	Reference	Reference
Two	3,446(92.9)	1.64(1.31-2.05)**	1.51(1.19-1.90)**	1,365(80.7)	1.94(1.62-2.33)**	1.38(1.14-1.68)**	2,056(97.0)	1.81(1.30-2.53)**	1.34(0.94-1.91)
Three or more	850(93.1)	1.66(1.14-2.52)**	1.59(1.06-2.37)**	226(86.2)	2.90(1.82-4.62)**	1.72(1.06-2.79)**	307(97.5)	2.21(0.88-5.57)	1.37(0.55-3.40)
Exposure to mass media									
Yes	926(83.1)	Reference	Reference	3,491(69.2)	Reference	Reference	4,215(94.4)	Reference	Reference
No	6,908(92.2)	2.39(1.85-3.08)**	1.49(1.15-1.95)**	1,101(85.0)	2.52(2.08-3.06)**	1.46(1.19-1.80)**	2,153(97.8)	2.64(1.85-3.79)**	1.65(1.13-2.40)**
Access to money for medication									
No big problem	2,271(88.4)	Reference	Reference	3,299(69.2)	Reference	Reference	2,992(94.2)	Reference	Reference
Big problem	5,573(92.1)	1.53(1.23-1.89)**	1.18(0.94-1.46)	1,291(82.1)	1.92(1.64-2.24)**	1.23(1.03-1.46)**	3,376(96.7)	1.79(1.35-2.39)**	1.24(0.90-1.71)
Travel distance for medication									
No big problem	4,722(90.3)	Reference	Reference	2,612(67.5)	Reference	Reference	3,467(94.8)	Reference	Reference
Big problem	3,121(92.0)	1.24(0.99-1.55)	0.92(0.73-1.15)	1,978(79.9)	2.04(1.71-2.44)**	1.25(1.02-1.52)**	2,902(96.4)	1.49(1.10-2.01)**	0.99(0.70-1.41)
Household head									
Male	6,268(90.9)	Reference	Reference	3,153(72.1)	Reference	Reference	5,180(95.4)	Reference	Reference

Female	1,578(91.4)	1.06(0.82-1.37)	1.09(0.82-1.45)	1,439(73.0)	1.05(0.89-1.23)	0.96(0.80-1.15)	1,189(96.0)	1.16(0.81-1.67)	1.06(0.0.72-1.56)
Household size									
4 or less	1,839(91.0)	Reference	Reference	1,673(64.0)	Reference	Reference	1,693(93.9)	Reference	Reference
5 or more	6,007(91.0)	1.00(0.78-1.29)	1.05(0.79-1.40)	2,919(78.3)	2.03(1.74-2.37)**	1.45(1.22-1.72)**	4,675(96.1)	1.60(1.20-2.15)**	1.18(0.86-1.61)
Household wealth									
Rich	3,198(94.3)	Reference	Reference	1,663(59.7)	Reference	Reference	2,422(91.9)	Reference	Reference
Middle	1,597(92.6)	1.88(1.37-1.59)**	1.05(0.76-1.44)	910(77.1)	2.27(1.85-2.78)**	1.49(1.20-1.86)**	1,219(96.9)	2.74(1.86-4.03)**	1.67(1.07-2.60)**
Poor	3,051(87.0)	2.48(1.94-3.18)**	0.94(0.66-1.35)	2,019(84.9)	3.78(3.18-4.48)**	1.89(1.53-2.33)**	2,728(98.4)	5.44(3.78-7.83)**	2.89(1.84-4.54)**

*n% is weighted counts and proportions** is a significant p-value<.05

6.2.3 Predictive modelling

Among all the Machine Learning prediction models, the Random Forest had better prediction performance based on the model prediction accuracies. The model performed at an accuracy of 75.7% implying that among the 30,404 instances (after using K-means SMOTE), the model correctly classified 23,016 instances. Of the total 23,016 instances, the model had Precision of 79.7%, Recall of 92.1%, Specificity of 51.6%, and an F1 score of 85.1% (Table 6.3). The AUROC was 70% (Figure 6.2). Household wealth, place of residence, maternal education and exposure to mass media were the top four most influential predictors of the drop out from the MNCH continuum (Figure 6.3).

Table 6.3: Prediction analysis of the drop out from the MNCH continuum using combined DHS surveys data for the DRC, Kenya and Tanzania (2013-2016)

Algorithm	Accuracy	Precision	Recall	Specificity	F score
Logistic regression	74.3	77.2	92.2	53.5	84.0
Random Forest	75.7	79.7	92.1	51.6	85.1
Decision tree	75.2	78.6	92.0	51.6	84.8
Support Vector Machine	75.3	78.8	91.9	51.0	84.8
Artificial Neural Network	75.4	78.5	92.3	53.3	84.8

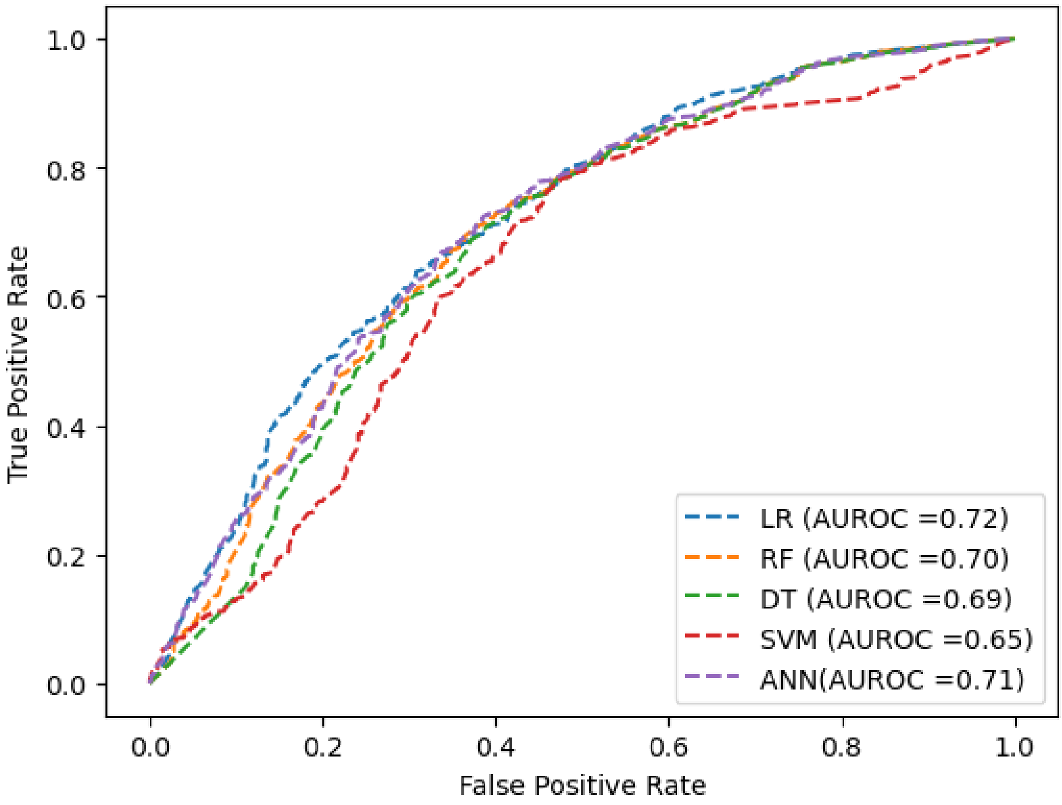

Figure 6.2: AUROC parameters for the five ML classification models

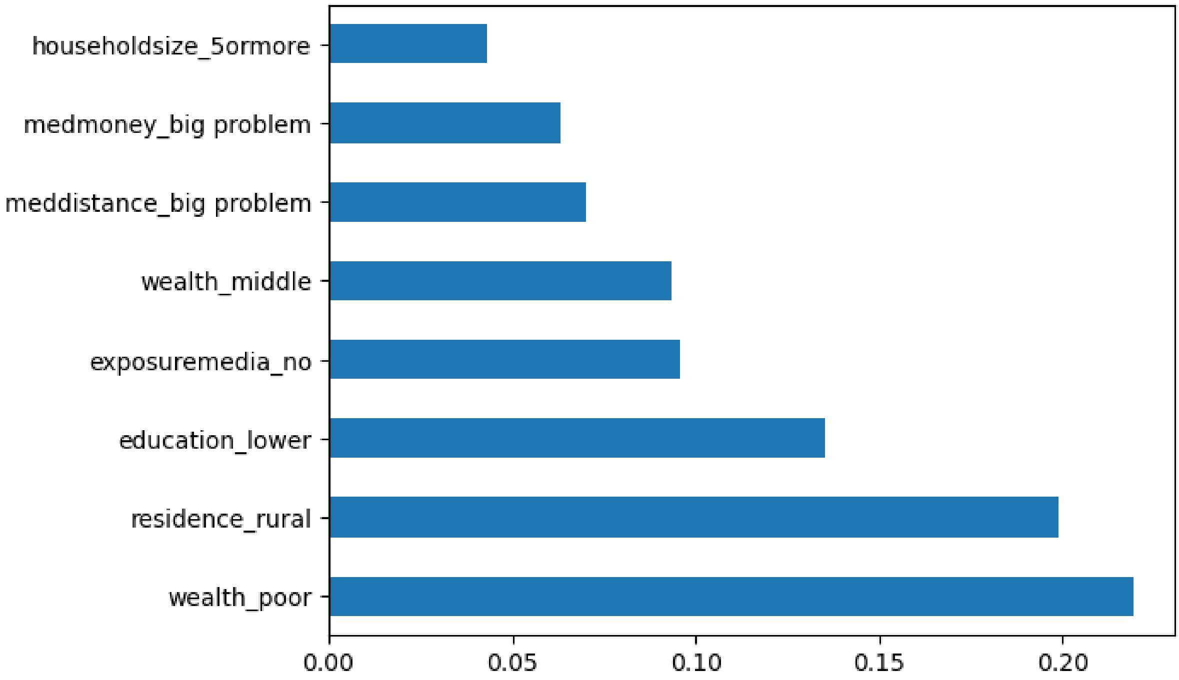

Figure 6.3: Feature (variable) ranking using the Random Forest

6.3 Discussion

The present study illustrated the determinants of the drop out from the MNCH continuum and developed Machine Learning models to forecast the drop out from the MNCH continuum using nationally representative DHS survey data from three sub-Saharan African countries including the DRC, Kenya, and Tanzania. The study findings showed that most women dropped out from the continuum of MNCH across the three countries. The prevalence of the drop out from the MNCH continuum was 91.0% in the DRC, 72.4% in Kenya and 93.6% in Tanzania. The largest gap and contributor to the high drop out from CoC was observed at PNC. The prevalence of the drop out from the MNCH continuum in the present study was consistent with other studies (7, 202). The high drop out in the MNCH continuum poses a higher risk of maternal and neonatal morbidity and mortality to many women and children due to missed opportunities for interventions in the CoC (197).

The study found that rural dwellers were significantly more likely to drop out from the MNCH continuum in the DRC, Kenya and Tanzania. These findings have been reported in other studies, where rural women were less likely to complete the CoC (7, 28, 203). These results are consistent with the view that rural women may encounter higher obstacles to obtaining maternal healthcare than urban women due to constrictive cultural norms, long travel distances, unaffordability of medicine, and the burden of caring for larger families (7, 197). Additionally, our analysis indicated that dropping out of the MNCH continuum, particularly in Kenya was linked with unaffordable medicine, long travel distances, and larger households. These factors may represent potential obstacles for rural Kenyan women's retention in the CoC (204, 205). Scaling up community education programs and developing policies that improve access to health facilities and supply and affordability of drugs could improve the retention of rural women in the continuum of MNCH (206).

The study revealed that low maternal education attainment was a positive predictor of dropping out from the MNCH continuum in the DRC, Kenya and Tanzania. Previous studies support our findings, the studies found that less educated women were more likely to be retained in maternity care (7, 31). Better education is believed to be an essential factor in creating better awareness and increasing knowledge of the importance of MNCH. Thus, government policies that promote female education and community interventions such as media campaigns would increase the awareness of women on the importance of the continuum of MNCH (93). The role of mass media was also evident in determining retention in MNCH as cited in previous research (178). Women who were not exposed to mass media were more likely to drop out of the MNCH continuum in the three countries. These results demonstrate the crucial role of the media in promoting and raising awareness about the continuation of MNCH. Thus, governments and other non-governmental organisations should invest continuously in the design and implementation of MNCH services utilisation educational programs through mass media channels to increase the use of these services (178).

The study also found that older women in the DRC and Kenya were less likely to drop out from the continuum of MNCH. However, there is a lack of consensus on the influence of the mother's age on maternal health utilisation (28, 207, 208). Our findings are similar to a previous study in Ghana which found that older women were less likely to drop out from maternal healthcare (207). This could be the reason that older women aged 25-49 years have gathered immense knowledge on the need to utilise maternal healthcare, which may positively influence their use of these services.

Higher birth order was positively associated with drop out from the MNCH in the DRC and Kenya. These findings corroborate results from other studies (7, 28). Possibly high parity women place high value on lower pregnancies because of experience in pregnancy and childbirth (209, 210). The retention of multigravida women in the continuum of MNCH could be enhanced by expanding community education programs through mass media campaigns.

Women in poorer households were more likely to drop out from the continuum of MNCH in Kenya and Tanzania. These findings were consistent with other studies in Kenya and elsewhere (28, 211, 212). Although MNCH services are being provided for free in Kenya and Tanzania, other factors, such as inadequate healthcare provision and transportation costs, may act as impediments to the complete utilisation of the MNCH continuum among the poor (212-215). The indirect costs for transportation, medication, and healthcare-related services might have contributed to the differences observed in the drop out from the CoC among the poor and rich women (212-215). A multi-pronged approach to addressing barriers to accessing care among the poor is required, taking into account other potential barriers such as travel costs to health facilities and a lack of staff or medication. (214).

Regarding the Machine Learning predictive analysis, this study showed that Machine Learning methods predict the drop out from the MNCH continuum better than the conventional logistic regression method. The Machine Learning model performance results showed that the logistic regression model had the lowest prediction accuracy compared to the other ML classification models. This result is not surprising, since Machine Learning methods are documented to outperform conventional logistic methods in several fields of medicine (34, 200). Our results also showed the Random Forest had the highest prediction accuracy compared to the rest of the models. These results showed that the Random Forest is the most suitable algorithm in this study to accurately predict the drop out from the MNCH continuum. The Random Forest is a commonly used Machine Learning model which combines the output of multiple decision trees to reach a single result. It is easily interpretable and flexible as compared to other Machine Learning algorithms such as Artificial Neural Network and Support Vector Machine (35).

Using the Random Forest, the study further ranked the most important predictors associated with the drop out from the MNCH continuum. Household wealth, place of residence, maternal education and exposure to mass media were the top four important predictors. The use of Machine Learning analysis can be valuable in identifying the most influential predictors for targeted interventions. This information can accelerate the improvement of the utilisation of the MNCH continuum in the SDG era, as it provides public health programmers and policymakers with cost-effective interventions for time and resource management (216). Rapid response mechanisms such as web-based applications can also be developed by applying Machine Learning. For instance, web-based applications can be used to assess the probability of a pregnant woman and unborn child dropping out from the CoC based on the mother's characteristics (34, 217). This allows for the provision of targeted interventions to pregnant women at high risk of discontinuing care in real-time and improves retention in the MNCH continuum (34, 217).

6.3.1 Strengths and Limitations

The study analysed factors contributing to the drop out from the MNCH continuum among several countries in the SSA. Thus, highlighting the common driving factors which should be considered when designing policies and interventions aimed at improving retention in the MNCH continuum. The study also developed Machine Learning models to forecast the drop out from the MNCH continuum, which is computationally strong when handling big data and can be used to classify certain hidden information that could not be detected by conventional statistical methods. However, the study is subject to several limitations. The main components of MNCH rely on the women's self-report which are subject to recall bias. Another possible limitation is that additional features that could have contributed to the prediction output were not present. This includes information that was not collected in the surveys such as health service provision features (quality of care and availability of drugs and equipment). The Machine Learning method is also novel in the SSA, we did not have enough evidence to compare the findings on the prediction of the drop out from MNCH continuum with other countries in SSA. The Machine Learning analysis did not account for survey design. This is because most Machine Learning methods were built for predictions and not ascertaining relationships. However, significant factors/features from the multivariable logistic regressions models with survey design correction were used in the Machine Learning predictions. Finally, it is important to state that both Conventional and Machine Learning techniques should be embraced as each method come with its own strengths which should be taken advantage of depending on the problem to be solved.

6.4 Conclusions

The prevalence of drop out from the MNCH continuum was 91.0% in the DRC, 72.3% in Kenya and 93.6% in Tanzania. The greatest contributor to the drop out from the continuum of MNCH was between delivery and the postpartum period. Place of residence, maternal education and exposure to mass media were common contributing factors associated with drop out from the MNCH continuum in the DRC, Kenya and Tanzania. Among the developed Machine Learning models, the Random Forest had better prediction accuracy. The top four predictors with the greatest influence were household wealth, place of residence, maternal education and exposure to mass media. The results of these findings can help inform evidence-based decisions in MNCH interventions and can also be used to assist in developing web-based applications that help public health practitioners take preventative action to retain more mothers and children in CoC.

6.5 Chapter Summary

Receiving the full package of antenatal, intrapartum, and postnatal care is necessary for the mother/child's optimal health. However, the drop out from the MNCH continuum remains a challenge in SSA. Few studies have assessed the drop out/full utilisation of the MNCH continuum. Using Machine Learning methods, this study sought to predict the MNCH continuum drop out and determine important predictors in three sub-Saharan African countries: the DRC, Kenya and Tanzania. The MNCH continuum dropout rate is very high in the DRC, Kenya and Tanzania. Maternal education, place of residence, and mass media exposure were common predictors of the drop out from MNCH continuum. Among the Machine Learning models, the Random Forest exhibited more accurate predictions of dropping out of the MNCH continuum. Household wealth, place of residence, maternal education and exposure to mass media were the top four predictors with significant influence. The findings of this study can be used in supporting evidence-based decisions in MNCH interventions and developing web-based services to improve retention of mothers and children in MNCH.

CHAPTER 7-EMPIRICAL EVIDENCE for objective 4

Mediation role of maternal, newborn and child healthcare services utilisation on neonatal mortality in three sub-Saharan African countries: A generalized structural equation modelling approach.

Authors: Chenai Mlandu[1], Zvifadzo Matsena-Zingoni[2], Eustasius Musenge[2]

Affiliation: School of Public Health, University of Witwatersrand, South Africa[1,2]

*chenai.mlandu@gmail.com

Corresponding author: Chenai Mlandu

(Under Review: Maternal Health, Neonatology and Perinatology)

Abstract

Background: Utilisation of maternal and newborn healthcare (MNCH) services has a bearing on neonatal health outcomes. An understanding of the interrelated pathways between MNCH services utilisation and neonatal health outcomes is crucial for more effective targeted interventions. Thus, this study sought to describe the mediation role of MNCH services utilisation on neonatal mortality in three sub-Saharan African countries.

Methods: This was a secondary analysis using cross-sectional Demographic Health Survey (DHS) data from the Democratic Republic of Congo (DRC) (2013/14), Kenya (2014) and Tanzania (2015/16) among women of reproductive age. The five interrelated (endogenous) outcomes were antenatal care (ANC) attendance, skilled birth attendance (SBA) and postnatal care (PNC) attendance, low birth weight and newborn mortality. Generalized Structural Equation Modelling (GSEM) was used to describe the relationships and considered cluster random effects.

Results: ANC attendance mediated the total effects of PNC attendance by 8.8% in Kenya and 5.5% in Tanzania. ANC attendance and SBA also sequentially mediated the total effects of

PNC attendance on neonatal mortality by 1.9% in Kenya and 1.0% in Tanzania. ANC attendance also mediated 2.2% of the total effects of LBW on neonatal mortality. In the DRC, no presence of mediation was observed, however, ANC attendance moderated the relationship between parity and neonatal mortality.

Conclusion: ANC attendance and SBA were identified as mediating factors of neonatal mortality. The mediation role of MNCH services on neonatal mortality indicates critical areas for intervention to reduce neonatal mortality. The findings highlighted the importance of targeted interventions among pregnant women and women who just delivered to seek maternal and child healthcare services from skilled health workers. Furthermore, low-cost interventions for LBW such as health education and home care practices in low-income settings should be strengthened.

Keywords; Maternal and Newborn Healthcare, Mediation, Neonatal mortality

7.0 Introduction

According to the World Health Organization (WHO), neonatal mortality is defined as the death of a child in the first month of life (128). The risk of a child's death within the first month of life is high and varies by region (160). A child born in Sub-Saharan Africa (SSA) is 9 times more likely to die within the first month of life compared to a child born in a high-income country (160). In 2017, SSA recorded the highest NMR at 27 deaths per 1000 live births, followed by Southern Asia with 26 deaths per 1,000 live births. Current projections show that many countries in SSA will fail to meet the SGD-3 goal of reducing the NMR to 12 deaths per 1,000 live births by 2030 and ending preventable neonatal mortality (2). In SSA, the DRC, Kenya and Tanzania are ranked in the top six with the highest neonatal deaths. In 2017, NMRs were estimated at 29 deaths per 1,000 births in 2017 in the DRC, and 21 deaths per 1,000 live births which are above the SDG- 3 target of 12 deaths per 1,000 live births by 2030 (2).

On a global scale, the main causes of neonatal mortality are prematurity (28%), sepsis (26%) and asphyxia/birth related hypoxia (23%) (17). In Africa neonatal deaths are mainly attributed to infections including, sepsis/pneumonia, tetanus, and diarrhoea, preterm birth, asphyxia and congenital defects (16, 218). However, most of the leading causes of neonatal deaths are preventable and treatable (17, 218). Several research has indicated numerous factors contributing to neonatal mortality (17, 50, 219-223). Factors found to be associated with neonatal mortality in developing countries include maternal age, mother's education, marital status, father's education, sex of the neonate, economic status, pregnancy complication, media access, birth weight of the neonate, home delivery, mode of delivery, lack of preparedness of families and care providers, harmful cultural practices, negative parental attitudes arising from the social environment, inability to pay for care, lack of prenatal and postnatal care (17, 36, 50, 219-223) .

An important aspect of preventing neonatal mortality is identifying modifiable factors linked to neonatal deaths. The high burden of neonatal mortality in SSA has been linked to several modifiable factors including ANC utilisation, SBA and PNC uptake (49, 220). However, utilisation of MNCH services i.e. ANC, SBA and PNC remains suboptimal in SSA (10, 224). Findings from a previous study found that optimal ANC utilisation i.e. attendance of four or more ANC visits was 55.5% (26). Another study found that the pooled prevalence of SBA was 63.9% (27) . PNC had been cited as the least utilised service in MNCH, with a prevalence below 54% (11). Furthermore, low uptake of a service at a particular stage in Continuum of Care (CoC) will have a negative impact on utilisation of subsequent services (65).

The associations between MNCH services utilisation and neonatal mortality have often been determined using approaches that investigate direct associations rather than indirect associations (17, 18, 36, 49, 225). In this study, we sought therefore to breach this gap by exploring the mediation mechanisms of MNCH services utilisation on neonatal mortality. The findings of our study can be utilized by health programmers, policymakers, and other stakeholders to effectively shape policy formulation for the development of interventions that aim to achieve the third SDG of 12 neonatal deaths per 1000 live births by 2030 (25). Furthermore, our research findings can enhance the current body of knowledge, specifically in comprehending the intricacies surrounding neonatal mortality and its associations with ANC, SBA, and PNC utilisation.

7.1 Methods

7.1.1 Study area and setting

The study utilised secondary data from DHS surveys in three sub-Saharan African countries including the DRC (2013/14), Kenya (2014), and Tanzania (2015/16) (95, 96). The three countries were selected for this study because they rank in the top six countries in SSA with the highest number of neonatal deaths (14).

7.1.2 Data source design and sampling

The DHS surveys are nationally representative cross-sectional household surveys. The DHS sample in the DRC, Kenya and Tanzania was stratified and selected in two stages. In the first stage, EAs or clusters were selected with probabilities proportional to EA size. In the second stage, households were selected per cluster with equal-probability systematic sampling. All women (residents and visitors) aged 15–49 years at a particular household were eligible to be interviewed (142). A total of 8,940, 6,596 and 6,873 women from the DRC, Kenya and Tanzania respectively were considered for the final analysis.

7.1.3 Inclusion and exclusion criteria

The study comprised women aged 15-49 years who had a live birth in the past 5 years preceding each country ‘s DHS survey. Only women who reported on ANC, SBA and PNC were included in the study. Only information related to the most recent birth was utilised to minimise measurement and recall bias.

7.1.4 Measurement and variables

Outcome and explanatory variables

Endogenous variables

The five endogenous variables were ANC attendance, coded as 1=low if ANC visits were less than 4, 0=high if otherwise, SBA, coded as 1=no if the mother was not attended by a skilled worker at delivery, and 0=yes if otherwise, PNC attendance, coded as 1=no if mother and child did not have PNC within six weeks of delivery and 0=yes if otherwise, LBW coded as 1=yes if the newborn had birth weight less than 2,500g and 0=no if otherwise, neonate mortality coded as 1=yes if neonate's death was within the first month of life, and 0=no if otherwise.

Exogenous variables

The study developed a Conceptual framework adapted from the Socio-Ecological Model (1) to select important predictors based on literature and describe potential pathways for interventions to reduce neonatal mortality, Figure 7.1. The Socio-Ecological Model is a comprehensive model of health that encompasses various levels of influence including individual, relationship, social, organisational and public policy factors. The individual factors included maternal and child characteristics such as maternal age (15-24 years, 25-49 years), maternal education (primary and no education, secondary and tertiary), wanted last pregnancy (then, later, no more), financial decision making (respondent, joint with husband/partner/husband, partner/other, has no partner), mode of delivery (normal, caesarean section), timing of ANC (early if gestational age is less than 4 months, late if gestational age is 4 or more months), employment in the past 12 months (no, yes) and child sex (male, female). Interpersonal factors included has a male partner (yes, no). Community factors comprised exposure to mass media (no, yes), household wealth (poor, middle, rich), household head (male, female), household size (4 or less, 5 or more) and permission to seek medical help alone (no big problem, big problem). Organisational factors comprised place of delivery (facility, home), travel distance to the health facility (no big problem, big problem), medical costs (no big problem, big problem), health insurance cover (yes, no) and quality of ANC. Quality of ANC was constructed using items on routine ANC services including measurement of weight, height and blood pressure, and collection of urine and blood samples. The items on routine ANC services were coded as binary responses (1=yes if the service was received and 0=no if otherwise). Women who scored >75% of the total score were classified as having received high quality of ANC, and low quality of ANC if otherwise.

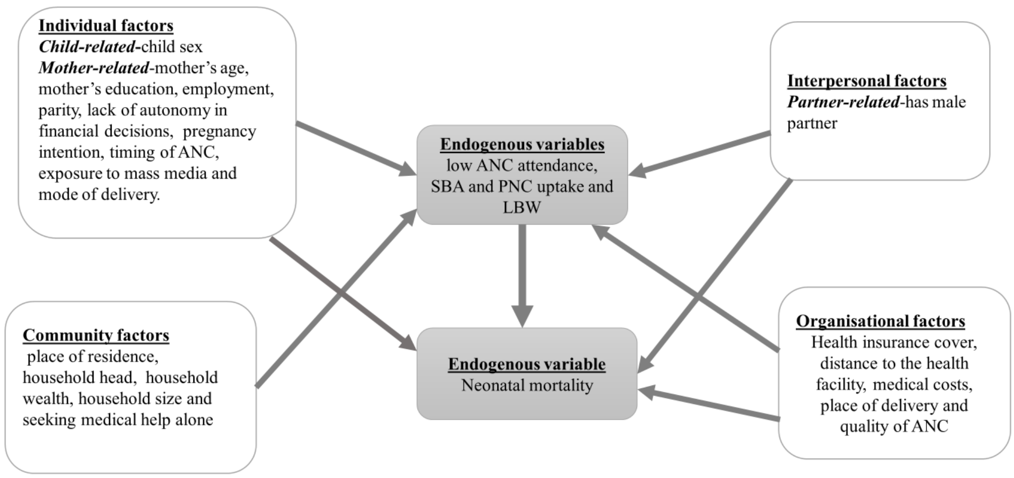

Figure 7.1: A conceptual framework for the mediation role of MNCH services utilisation on neonatal mortality [adapted from the Socio-Ecological Model (1)].

7.1.5 Data analysis

Data was analysed using Stata version 17. All analysis considered the survey design by accounting for clustering, survey weights and stratification using the *svyset* command. A descriptive summary of the participants' characteristics was done using Pearson's chi square test. To examine the MNCH services utilisation and newborn mortality pathways, several analysis steps were applied. Multivariate analysis with all covariates were performed to identify the variables that would significantly affect outcomes (endogenous variables) of interest in the adjusted models. The results were reported in terms of ORs and 95% CI. The level of significance was determined at 5%.

The analysis ran GSEM using explanatory variables that were significant in the multivariable logistic regression analysis. GSEM evaluates potential causal relationships with the "structural model" and considers both direct and indirect effects of multiple interacting factors. In GSEM, it is possible to construct a model with both continuous and discrete variables grouped together in the same latent construct (130). Therefore, GSEM was with binomial logit link function (130) was used to test for the mediation role of MNCH services utilisation on newborn mortality. The GSEM analysis accounted for survey weights and incorporated random effects at cluster level. The random effects show the variation of each cluster/community on neonatal mortality. The Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) were used to assess if adding paths in the GSEM model improved the model and whether the final selected model fit with random effects at the cluster level had a better fit than the model without random effects. A model with a lower BIC was considered a model with the best fit (132).

The mediation analysis was performed using Stata *'gsem'* command. A mediator is a variable that explains the relationship between an outcome and explanatory variable. Mediation can be partial or complete mediation (133, 134). Complete mediation is present when the exposure variable no longer influences the outcome variable after the mediator has been controlled (133, 134). Partial mediation occurs when the explanatory variable's influence on the outcome variable is reduced after the mediator is controlled (133, 134). Mediators can also be classified as single and sequential. A single mediator refers to when there is only one variable in the causal pathway between exposure and outcome variable. Multiple mediators are considered when more than one mediator variables operate jointly at the same stage in a causal model (135). Thus, there will be several indirect effects linking the exposure variable to the outcome variable.

Stata *'nlcom'* command was used to estimate the direct, indirect, and total effects of late ANC initiation on underweight. The indirect effects were estimated using the product of the coefficients test. Stata *'nlcom'* command was used to estimate the direct, indirect and total effects.

7.2 Results

7.2.1 Characteristics of study participants

Table 7.1 presents the characteristics of study participants in the DRC (8,940), Kenya (6,596) and Tanzania (6,873). A greater proportion of the women lived in the rural areas in the DRC (64.3%), Kenya (60.1%) and Tanzania (69.7%). Majority of the women were aged between 25-49 years in the DRC (69.4%), Kenya (70.8%) and Tanzania (67.3%). A greater proportion of the women in the DRC (56.6%), Kenya (63.1%) and Tanzania (83.6%) had primary or no education. Women of three or more parity were 11% in the DRC (10.8%), 4% in Kenya (4.2%) and 5% in Tanzania (4.7%). The proportion of female newborns was 50% in the DRC (50.2%), 49% in Kenya (48.7%) and Tanzania (48.6%). Women with unwanted last pregnancies were in 6% in the DRC (6.0%), 12% in Kenya (11.7%) and 5% in Tanzania (4.7%). Unemployed women in the past 12 months were about 20% in the DRC (20.3%), 28% in Kenya (28.4%) and 16% in Kenya (16.2%).

Women without male partners were 15% in the DRC (15.3%), 18% in Kenya (18.3%) and 20% in Tanzania (19.7%). Most of the households were led by males in the DRC (79.7%), Kenya (69.0%) and Tanzania (81.3%). Women from poor-wealth households were 39% in the DRC

(39.3%), 38% in Kenya (37.5%) and 42% in Tanzania (41.2%). Majority of the women lived in households of 5 or more members in the DRC (75.9), Kenya (58.3) and Tanzania (72.4). Women with no media access were 54% in the DRC (54.2%), 14% in Kenya (14.0%) and 17% in Tanzania (17.1%). Women with husbands/partners/others who made financial decisions for them were 35% in the DRC (35.1%), 36% in Kenya (35.6%) and 33% Tanzania (33.3%).

Women who had big problems with seeking medical help alone in the DRC were about 33% in the DRC (33.0%), 6% in Kenya (6.2%) and 14% in Tanzania (13.9%). About 39% of the women in the DRC (39.2%), 25% in Kenya (24.9%) and 45% in Tanzania (45.1%) had big problems with travel distance to health facilities. Nearly 70% of the women in the DRC (70.4%), 39% in Kenya (39.1%) and 52% in Tanzania (52.3%) had big problems with medical costs. A greater proportion of women in the DRC (95.9%), Kenya (81.9%) and Tanzania (92.3%) did not have health insurance cover. Women who initiated ANC late were 61% in the DRC (60.5%), 58% in Kenya (58.2%) and 50% in Tanzania (49.6%). Nearly 47% of women in the DRC (46.8%), 39% in Kenya (39.4%) and 48% in Tanzania (47.9%) had low ANC attendance.

The proportion of women who received low quality ANC were 48% in the DRC (47.5%), 42% in Kenya (42.0%) and 47% in Tanzania (47.2%). A greater proportion of women in the DRC (89.3%), Kenya (84.5%) and Tanzania (66.3%) had no skilled attendance at delivery. About 12% of women in the DRC (12.0%), 30% in Kenya (30.4%) and 31% in Tanzania (31.1%) delivered at home. Nearly 6% of the women in the DRC (6.3%), 10% in Kenya (9.5%) and 7% in Tanzania (7.0%) delivered by caesarean section. A greater proportion of women and their children in the DRC (84.6%), Kenya (53.8%) and Tanzania (79.1%) did not utilise PNC. Newborns with low birth weight were 6% in the DRC (6.4%), 7% in Kenya (6.9%) and 6% in Tanzania (6.2%). Neonatal mortality was 2% in the DRC (2.1%), 1% in Kenya (1.4%) and 2% in Tanzania (1.8%).

Table 7.1: Characteristics of study participants in the DRC, Kenya and Tanzania

	DRC N(8,940) * n(%)	Kenya N(6,596) * n(%)	Tanzania N(6,873) *n(%)
Place of residence			
Urban	3,212(35.7)	2,602(39.9)	2,083(30.3)
Rural	5,788(64.3)	3,910(60.1)	4,798(69.7)
Maternal age			
Young women (18-24 years)	2,756(30.6)	1,901(29.2)	2,249(32.7)
Older women (25-49 years)	6,244(69.4)	4,612(70.8)	4,632(67.3)
Maternal education			
Primary and below	5,086(56.6)	4,109(63.1)	5,754(83.6)
Secondary and above	3,914(43.5)	2,402(36.9)	1,127(16.4)
Employment in the past 12 months			
No	1,824(20.3)	1,851(28.4)	1,111(16.2)
Yes	7,175(79.7)	4,659(71.6)	5,770(83.8)
Parity			
One	4,188(46.5)	4,495(69.0)	4,380(63.7)
Two	3,842(42.7)	1,744(26.8)	2,176(31.6)
Three or more	971(10.8)	273(4.2)	325(4.7)
Child sex			
Male	4,482(49.8)	3,338(51.3)	3,536(51.4)
Female	4,518(50.2)	3,175(48.7)	3,345(48.6)
Wanted last pregnancy			
Then	6,028(67.0)	3,933(60.4)	4,558(66.3)
Late	2,425(27.0)	1,818(27.9)	1,983(28.8)
No more	546(6.0)	761(11.7)	340(4.9)
Has a male partner			
No	1,379(15.3)	1,194(18.3)	1,354(19.7)
Yes	7,621(84.7)	5,320(81.7)	5,592(80.3)
Household head			
Male	7,174(79.7)	4,491(69.0)	5,592(81.3)
Female	1,826(20.3)	2,023(31.0)	1,289(18.7)
Household wealth			
Poor	3,533(39.3)	2,443(37.5)	2,837(41.2)
Middle	1,821(20.2)	1,218(18.7)	1,316(19.1)
Rich	3,646(40.5)	2,850(43.8)	2,727(39.7)
Household size			
4 or less	2,172(24.1)	3,714(41.7)	1,898(27.6)
5 or more	6,828(75.9)	3,800(58.3)	4,983(72.4)
Exposure to mass media			
No	4,874(54.2)	914(14.0)	1,175(17.1)
Yes	4,126(45.8)	5,597(86.0)	5,705(82.9)
Financial decisions			
Respondent alone	687(7.7)	441(6.9)	242(3.5)
Joint with husband/partner	3,692(41.6)	2,494(38.9)	2,971(43.4)
Husband/partner/other	3,199(35.1)	2,280(35.6)	2,278(33.3)
Has no partner	1,379(15.6)	1,194(18.6)	1,354(19.8)
Seeking medical help alone			
No big problem	6,028(67.0)	6,106(93.8)	5,921(86.1)
Big problem	2,967(33.0)	403(6.2)	960(13.9)
Travel distance to the health facility			
No big problem	5,471(60.8)	4,892(75.1)	3,776(54.9)
Big problem	3,524(39.2)	1,619(24.9)	3,105(45.1)
Medical costs			

No big problem	2,667(29.6)	3,963(60.9)	3,283(47.7)
Big Problem	6,330(70.4)	2,548(39.1)	3,598(52.3)
Health insurance cover			
No	8,632(95.9)	5,336(81.9)	6,351(92.3)
Yes	368(4.1)	1,176(18.1)	530(7.7)
Timing of ANC			
Early	3,557(39.5)	2,723(41.8)	3,470(50.4)
Late	5,443(60.5)	3,790(58.2)	3,410(49.6)
ANC attendance			
High	4,191(46.8)	2,545(39.2)	3,283(47.9)
Low	4,765(53.2)	3,946(60.8)	3,565(52.1)
Quality of ANC			
Low	4,278(47.5)	3,775(58.0)	3,251(47.2)
High	4,723(52.5)	2,738(42.0)	3,630(52.8)
SBA			
Yes	8034(89.3)	5,502(84.5)	4,564(66.3)
No	967(10.7)	1,009(15.5)	2,316(33.7)
Place of delivery			
Facility	7,920(12.0)	4,530(69.6)	4,742(68.9)
Home	1,080(88.0)	1,981(30.4)	2,139(31.1)
Mode of delivery			
Normal	8,424(93.7)	5,886(90.5)	6,401(93.0)
Caesarean section	566(6.3)	620(9.5)	480(7.0)
PNC attendance			
Yes	1,381(15.4)	3,007(46.2)	5,442(20.9)
No	7,619(84.6)	3,506(53.8	1,439(79.1)
Low birth weight			
No	6,942(93.6)	4438(93.1)	4422(93.8)
Yes	474(6.4)	329(6.9)	293(6.2)
Neonatal mortality			
No	8,812(97.9)	6,422(98.6)	6,759(98.2)
Yes	189(2.1)	92(1.4)	122(1.8)

7.2.2 The interrelationships between MNCH services and neonatal mortality

Table 7.2 indicates the details of the interrelationships between MNCH services utilisation and neonatal mortality in the DRC. The predictors of low ANC attendance were timing of ANC, place of residence, maternal education, birth order, medical costs, seeking medical help alone and quality of ANC. Women who initiated ANC late (aOR =6.59, 95% CI 5.18–8.39), had primary or no education (aOR =1.49, 95% CI 1.18–1.88), lived in the rural areas (aOR =1.48, 95% CI 1.15–1.90), had parity of two (aOR =1.24, 95% CI 1.08–1.43) and three or more (aOR =1.40, 95% CI 1.14–1.70) were significantly associated with low ANC attendance. Women who faced big problems of medical costs (aOR =1.33, 95% CI 1.11–1.60) and seeking medical help alone (aOR =1.28, 95% CI 1.08–1.52), and received low quality of ANC (aOR =1.38, 95% CI 1.06–1.79) were also significantly more likely to have low ANC attendance.

Table 7.2: Analysis of the interrelationships between MNCH services utilisation and neonatal mortality using GSEM in the DRC

	Low ANC attendance OR (95% CI)	SBA OR (95% CI)	PNC attendance OR (95% CI)	LBW OR (95% CI)	Neonatal mortality OR (95% CI)
Timing of ANC					
Early	[a]				
Late	6.59(5.18-8.39)	-	-	-	-
Place of residence					
Urban	[a]	[a]			
Rural	1.48(1.15-1.90)	2.27(1.38-3.75)	-	-	-
Maternal education					
Secondary and above	[a]		[a]		
Primary and below	1.49(1.18-1.88)	-	1.74(1.36-2.24)	-	-
Household wealth					
Rich		[a]			
Middle	-	2.40(1.55-3.71)	-	-	-
Poor	-	3.40(2.16-5.35)	-	-	-
Has male partner					
No				[a]	
Yes	-	-	-	0.72(0.52-0.99)	-
Parity					
One	[a]				[a]
Two	1.24(1.08-1.43)	-	-	-	0.88(0.53-1.45)
Three or more	1.40(1.14-1.70)	-	-	-	2.60(1.45-4.66)
Exposure to mass media					
Yes		[a]			

No	-	1.35(1.06-1.71)	-	-	-
Quality of ANC					
High	a	a	a		
Low	1.38(1.06-1.79)	1.54(1.12-2.11)	1.80(1.38-2.34)	-	-
Health insurance cover					
Yes		a	a		
No	-	3.10(1.06-9.04)	1.61(1.06-2.46)	-	-
Medical costs					
No problem	a			a	
Big problem	1.33(1.11-1.60)	-	-	1.47(1.08-2.01)	-
Travel distance to the health facility					
No problem		a			
Big problem	-	1.52(1.11-2.08)	-	-	-
Seeking medical help alone					
No big problem	a	a			
Big problem	1.28(1.08-1.52)	1.38(1.01-1.91)	-	-	-
Low ANC attendance					
Four or more		a			
less than four	-	1.55(1.22-1.97)	-	-	-
Place of delivery					
Facility			a		
Home	-	-	2.29(1.48-3.54)	-	-
Mode of delivery					
Normal			a		
Caesarean section	-	-	0.63(0.46-0.85)	-	-
LBW					
Yes					a
No	-	-	-	-	3.08(1.67-5.69)

OR-Odds Ratio, * $p < 0.05$, [a]reference category, dash (−) denotes variables not considered for the particular outcome

The predictors of SBA were place of residence, household wealth, exposure to mass media, quality of ANC, health insurance cover, travel distance to health facilities, seeking medical help alone and low ANC attendance. The odds of not having SBA were significantly higher among women who lived in the rural areas (aOR =2.27, 95% CI 1.38–3.75), had middle (aOR =2.40, 95% CI 1.55–3.71) and poor household wealth (aOR =3.40, 95% CI 2.16–5.35), had no access to mass media (aOR =1.35, 95% CI 1.06–1.71) and had no health insurance cover (aOR =3.10, 95% CI 1.06–9.04). The odds of not having SBA were also significantly higher among women who had big problems of travel distance to the health facility (aOR =1.52, 95% CI 1.11–2.08) and seeking medical help alone (aOR =1.38, 95% CI 1.01–1.91), had low ANC attendance (aOR =1.55, 95% CI 1.22–1.97) and received low quality of ANC (aOR =1.54, 95% CI 1.12–2.11).

The predictors of PNC attendance were maternal education, quality of ANC, health insurance cover, mode of delivery and place of delivery. The odds of not attending PNC were significantly higher among women who had primary or no education (aOR =1.74, 95% CI 1.36–2.24), low quality of ANC (aOR =1.80, 95% CI 1.38–2.34), had no health insurance cover (aOR =1.61, 95% CI 1.06–2.46) and delivered at home (aOR =2.29, 95% CI 1.38–2.34). However, women who delivered by caesarean section (aOR =0.63, 95% CI 0.46–0.85) had significantly lower odds of not attending PNC.

The predictors of low birth weight were medicals costs and relationship status. Women with big problems of medical costs (aOR =1.47, 95% CI 1.08–2.01) were significantly more likely to have LBW newborns, while women with male partners (aOR =0.72, 95% CI 0.52–0.99) were significantly less likely to have LBW newborns. The predictors of neonatal mortality were LBW and birth order. The odds of neonatal mortality were significantly higher among LBW new-borns (aOR =3.08, 95% CI 1.67–5.69) and women of parity of three or more (aOR =2.60, 95% CI 1.45–4.66).

Table 7.3 indicates the details of the interrelationships between MNCH services utilisation and neonatal mortality in Kenya. The predictors of low ANC attendance were timing of ANC, maternal education, parity, household wealth, seeking medical help alone, health insurance cover and quality of ANC. The odds of low ANC attendance were significantly higher among women who initiated ANC late (aOR =12.29, 95% CI 10.30–14.66), had primary or no education (aOR =1.37, 95% CI 1.09–1.73), had parity of two (aOR =1.39, 95% CI 1.19–1.64) and three or more (aOR =1.82, 95% CI 1.28–2.59), had middle (aOR =1.46, 95% CI 1.16–1.85) and poor (aOR =1.48, 95% CI 1.22–1.80) household wealth, had big problems of seeking medical help alone (aOR =1.34, 95% CI 1.04–1.72), had no health insurance cover (aOR =1.38, 95% CI 1.08–1.76) and received low quality of ANC (aOR =1.32, 95% CI 1.12–1.56).

The predictors of SBA were place of residence, household wealth, household size, maternal age, quality of ANC, health insurance cover and low ANC attendance. The odds of not having SBA was significantly higher among women who lived in the rural areas (aOR =1.45, 95% CI 1.13–1.98), had middle (aOR =2.21, 95% CI 1.54–3.18) and poor household wealth (aOR =3.88, 95% CI 2.88–5.22), lived in households with 5 or more members ((aOR =1.49, 95% CI 1.21–1.84), aged 25-49 years (aOR =1.60, 95% CI 1.32–1.93), had no health insurance cover (aOR =2.07, 95% CI 1.44–2.97), had low ANC attendance (aOR =1.51, 95% CI 1.28–1.77) and received low quality of ANC (aOR =1.31, 95% CI 1.08–1.59).

Table 7.3: Analysis of the interrelationships between MNCH services utilisation and neonatal mortality using GSEM in Kenya

	Low ANC attendance OR (95% CI)	SBA OR (95% CI)	PNC attendance OR (95% CI)	LBW OR (95% CI)	Neonatal mortality OR (95% CI)
Timing of ANC					
Early	[a]				
Late	12.29(10.30-14.66)	-	-	-	-
Place of residence					
Urban		[a]			
Rural	-	1.53(1.18-1.97)	-	-	-
Maternal education					
Secondary and above	[a]				
Primary and below	1.37(1.09-1.73)	-	-	-	-
Maternal age					
15-24 years		[a]			
25-49 years	-	1.76(1.46-2.13)	-	-	-
Household wealth					
Rich	[a]	[a]			
Middle	1.46(1.16-1.85)	2.30(1.59-3.33)	-	-	-
Poor	1.48(1.22-1.80)	4.04(2.99-5.48)	-	-	-
Household size					
4 or less		[a]			
5 or more	-	1.49(1.21-1.84)	-	-	-
Employed in the past year					
Yes			[a]		
No	-	-	1.50(1.27-1.76)	-	-
Parity					
One	[a]			[a]	
Two	1.39(1.19-1.64)	-	-	1.20(0.83-1.76)	-
Three or more	1.82(1.28-2.59)	-	-	2.03(1.09-3.80)	-
Exposure to mass media					
Yes			[a]		

No	-	-	1.30(1.08-1.58)	-	-
Quality of ANC					
High	a	a	a		
Low	1.32(1.22-1.56)	1.32(1.09-1.61)	1.23(1.06-1.43)	-	-
Health insurance cover					
Yes	a		a		
No	1.38(1.08-1.76)	-	1.35(1.11-1.65)	-	-
Medical costs					
No problem			a		
Big problem	-	-	1.23(1.06-1.43)	-	-
Seeking medical help alone					
No big problem	a		a		a
Big problem	1.34(1.04-1.72)	-	2.50(1.27-1.76)	-	3.45(1.40-8.53)
Low ANC attendance					
Yes		a	a		
No	-	1.53(1.30-1.80)	1.24(1.07-1.44)	-	-
Place of delivery					
Facility			a		
Home	-	-	3.10(3.00-3.69)	-	-
Mode of Delivery					
Normal			a		
Caesarean section	-	-	0.52(0.39-0.70)		
SBA				-	-
Yes			a		
No	-	-	1.86(1.46-2.37)	-	-
LBW					
No					a
Yes	-	-	-	-	2.280(0.99-5.25)
Child sex					
Male				a	
Female	-	-	-	1.43(1.07-1.92)	

OR-Odds Ratio, * $p < 0.05$, [a]reference category, dash (−) denotes variables not considered for the particular outcome

The predictors of PNC attendance were employment status, exposure to mass media, health insurance cover, ANC attendance, quality of ANC, medical costs, seeking medical help alone, SBA, place of delivery, health insurance cover, place of delivery and mode of delivery. The odds of not attending PNC were significantly higher among women who were unemployed (aOR =1.50, 95% CI 1.27–1.76), had no access to mass media (aOR =1.30, 95% CI 1.08–1.58), had no health insurance cover (aOR =1.35, 95% CI 1.11–1.65), had big problems of medical costs (aOR =1.23, 95% CI 1.06–1.43) and seeking medical help alone (aOR =2.50, 95% CI 1.86–3.37), delivered at home (aOR =3.10, 95% CI 3.00–3.69), had low ANC attendance (aOR =1.24, 95% CI 1.07–1.44), and received low quality of ANC (aOR =1.23, 95% CI 1.06–1.43). However, women who delivered by caesarean section (aOR =0.52, 95% CI 0.39–0.70) had significantly lower odds of not attending PNC.

Parity and child sex predicted LBW among newborns. Women of three or more parity had significant increased odds of having LBW newborns (aOR =2.04, 95% CI 1.09–3.80). Female newborns had significant odds of having LBW (aOR =1.43, 95% CI 1.07–1.92). The predictor of neonatal mortality was seeking medical help alone. The odds of neonatal mortality was significantly higher among women who had big problems of seeking medical help alone (aOR =3.45, 95% CI 1.40–8.53).

Table 7.4 indicates the details of the interrelationships between MNCH services and neonatal mortality in Tanzania. The predictors of low ANC attendance were timing of ANC, maternal education, parity, household wealth and quality of ANC.

Table 7.4: Analysis of interrelationships between MNCH services utilisation and neonatal mortality using GSEM in Tanzania

	Low ANC attendance OR (95% CI)	SBA OR (95% CI)	PNC attendance OR (95% CI)	LBW OR (95% CI)	Neonatal mortality OR (95% CI)
Timing of ANC					
Early	[a]				
Late	6.83(6.00-7.77)	-	-	-	-
Place of residence					
Urban		[a]			
Rural	-	2.03(1.27-3.24)	-	-	-
Maternal education					
Secondary and above	[a]	[a]			
Primary and below	2.78(1.94-3.93)	2.73(1.90-3.93)	-	-	-
Maternal age					
15-24 years		[a]			
25-49 years	-	1.21(1.03-1.42)	-	-	-
Household wealth					
Rich	[a]	[a]			
Middle	1.53(1.23-1.92)	1.89(1.42-2.53)	-	-	-
Poor	1.59(1.32-1.92)	2.99(2.26-3.95)	-	-	-
Household size					
4 or less		[a]	[a]		
5 or more	-	1.38(1.11-1.66)	1.24(1.06-1.45)	-	-
Wanted last pregnancy					
Then			[a]		
Late	-	-	1.07(0.91-1.26)	-	-
No more	-	-	1.63(1.08-2.46)	-	-
Has male partner					
No				[a]	[a]
Yes	-	-	-	0.56(0.43-0.73)	0.51(0.28-0.91)
Parity					
One	[a]				

Two	1.25(1.09-1.45)	-	-	-	-
Three or more	1.36(1.00-1.85)	-	-	-	-
Exposure to mass media					
Yes		[a]	[a]		
No	-	1.52(1.19-1.94)	1.39(1.11-1.74)	-	-
Quality of ANC					
High	[a]	[a]	[a]		
Low	1.66(1.44-1.92)	1.61(1.38-1.89)	1.62(1.37-1.92)	-	-
Health insurance cover					
Yes		[a]			
No	-	1.89(1.38-2.53)	-	-	-
Travel distance to the health facility					
No problem		[a]			
Big problem	-	1.32(1.13-1.54)	-	-	-
Seeking medical help alone					
No big problem			[a]		
Big problem	-	-	1.70(1.31-2.20)	-	-
Low ANC attendance					
Yes		[a]		[a]	
No	-	1.40(1.20-1.62)	-	1.55(1.19-2.01)	-
Place of delivery					
Facility			[a]		
Home	-	-	2.93(2.36-3.63)	-	-
Mode of delivery					
Normal			[a]		[a]
Caesar	-	-	0.43(0.35-0.54)	-	2.61(1.28-5.33)
PNC attendance					
Yes					[a]
No	-	-	-	-	6.74(2.23-20.38)
LBW					
Yes					[a]
No	-	-	-	-	3.56(1.75-7.24)
Child sex					
Male					[a]
Female	-	-	-	-	0.57(0.33-0.97)

Odds Ratio, * $p < 0.05$, [a]reference category, dash (−) denotes variables not considered for the particular outcome

The odds of having low ANC attendance were significantly higher among women who started ANC late (aOR =6.83, 95% CI 6.00–7.77), had primary or lower education (aOR =1.46, 95% CI 1.14–1.86), had parity of two (aOR =1.25, 95% CI 1.09–1.45), had middle (aOR =1.53, 95% CI 1.23–1.92) and poor (aOR =1.59, 95% CI 1.32–1.92) household wealth, and received low quality of ANC (aOR =1.66, 95% CI 1.44–1.92).

The predictors of SBA were place of residence, exposure to mass media, maternal education, maternal age, household wealth, household size, travel distance to the health facility, quality of ANC, health insurance cover and low ANC attendance. The odds of not having SBA was significantly higher among women who lived in the rural areas (aOR =2.03, 95% CI 1.27–3.24), had no access to mass media (aOR =1.52, 95% CI 1.19–1.94), had primary or no education (aOR =2.73, 95% CI 1.90–3.93), aged 25-49 years (aOR =1.21, 95% CI 1.03–1.42), had middle (aOR =1.89, 95% CI 1.42–2.53) and poor (aOR =2.99, 95% CI 2.26–3.95) household wealth. The odds of not having SBA was also significantly higher among women who lived in households of 5 or more members (aOR =1.38, 95% CI 1.11–1.66), had big problems with travel distance to the health facility (aOR =1.32, 95% CI 1.13–1.54), had no health insurance cover (aOR =1.89, 95% CI 1.38–2.58), had low ANC attendance (aOR =1.40, 95% CI 1.20–1.62) and received low quality of ANC (aOR =1.61, 95% CI 1.38–1.89).

The predictors of PNC attendance were exposure to media, wanted pregnancy, seeking medical help alone, place of delivery, mode of delivery, exposure to mass media, low ANC attendance and quality of ANC. The odds of not attending PNC were significantly higher among women who had no access to mass media (aOR =1.39, 95% CI 1.11–1.74), wanted pregnancy no more (aOR =1.63, 95% CI 1.08–2.46), lived in households of 5 or more members (aOR =1.24, 95% CI 1.06–1.45), had big problems of seeking medical help alone (aOR =1.70, 95% CI 1.31–2.20), delivered at home (aOR =2.93, 95% CI 2.36–3.63) and had low quality of ANC (aOR

=1.62, 95% CI 1.37–1.92). However, women who delivered by caesarean section (aOR =0.43, 95% CI 0.35–0.54) had significantly lower odds of not attending PNC.

Low ANC attendance and maternal age predicted LBW among newborns. The odds of LBW newborns was significantly higher among women with low ANC attendance (aOR =1.55, 95% CI 1.19–2.01), and lower among women aged 25-49 years (aOR =0.56, 95% CI 0.43–0.73). The predictors of neonatal mortality were LBW, relationships status, mode of delivery, PNC attendance and child sex. The odds of neonatal mortality was significantly higher among women who had LBW newborns (aOR =3.56, 95% CI 1.75–7.24), delivered by caesarean section (aOR =2.61, 95% CI 1.28–5.33) and had no PNC (aOR =6.72, 95% CI 2.23–20.38), and lower among women with male partners (aOR =0.51, 95% CI 0.28–0.91) and female newborns (aOR =0.57, 95% CI 0.33–0.97).

7.2.3 Mediation analysis

Table 7.5, Figure 7.2-4, shows the results of MNCH services mediation analysis on neonatal mortality. The results showed no presence of mediation, however, ANC attendance moderated the relationship between parity and neonatal mortality. Mothers of 3 or more parity were likely to have low ANC attendance and have deaths of their newborns. The direct determinants of ANC attendance as a moderating factor were timing of ANC, rural residence, quality of ANC, and health insurance cover, see Fig 7.2.

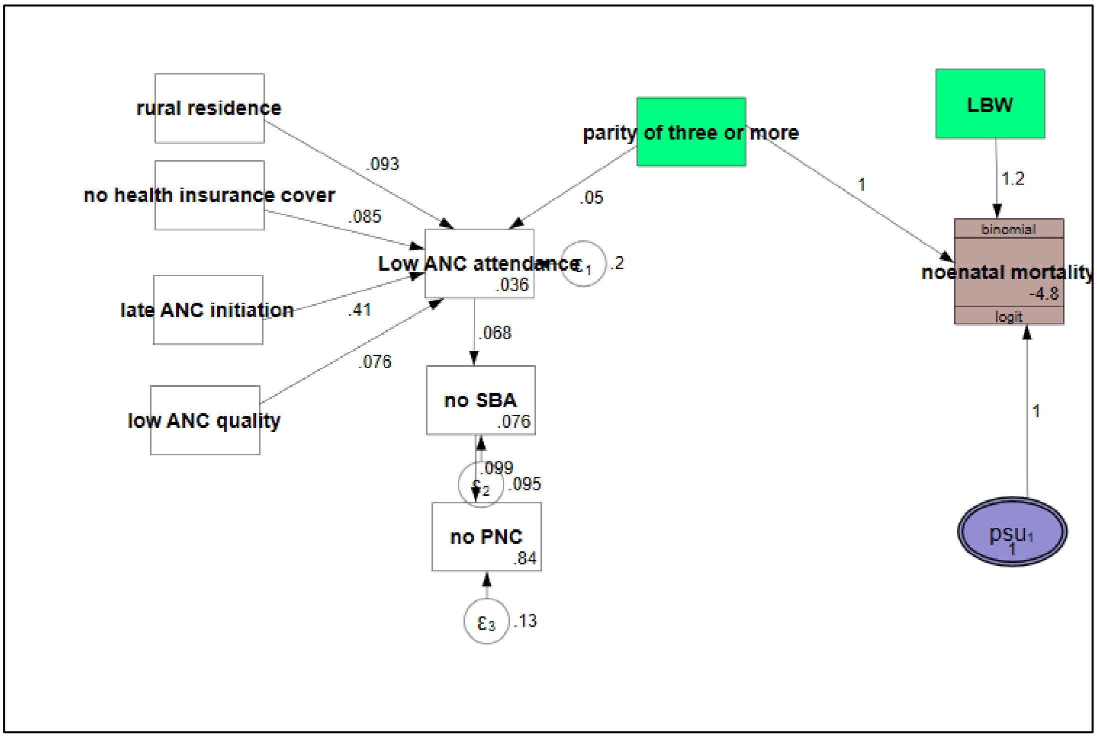

Figure 7.2: GSEM model of mediating role of MNCH services utilisation on neonatal mortality in the DRC. Key: variable type- light blue box (mediator), rose box (neonatal mortalty), mint box (direct predictors of neonatal mortality), white box (covariates) and lavender box (cluster random variable)

***Coefficients(β), all pathways are significant i.e. p<0.05**

In Kenya, ANC attendance and SBA were considered as mediating factors for neonatal mortality. ANC attendance mediated 8.8% of the total effects of PNC attendance on mortality. Women who had low ANC attendance were more likely not to have PNC and experience newborn deaths. The direct determinants of ANC attendance as a mediating factor were timing of ANC, quality of ANC, health insurance, place of residence and maternal education. ANC attendance and SBA sequentially mediated 1.9% of the total effects of PNC attendance on neonatal mortality. Women who had low ANC attendance were more likely not to have SBA, and women who did not have SBA were also more likely not to attend PNC and experience newborn deaths. The direct determinants of SBA as a mediating factor were place of residence, maternal education, health insurance and quality of ANC, see Figure 3.

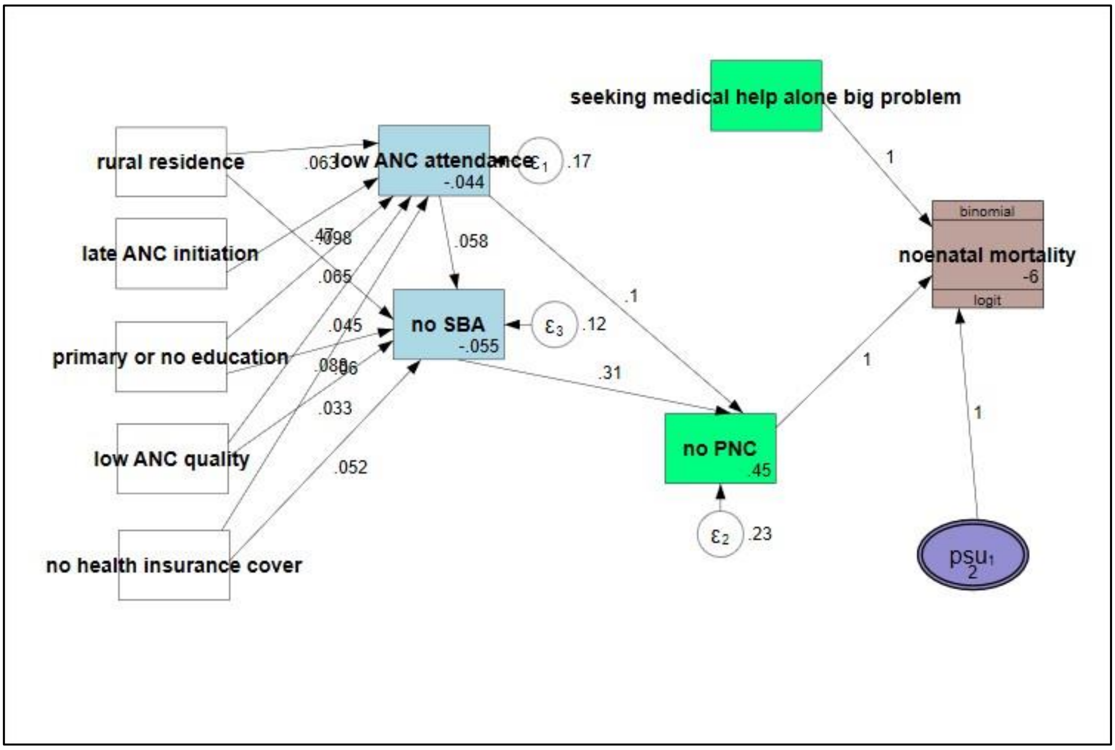

Figure 7.3: GSEM model of mediating role of MNCH services utilisation on neonatal mortality in Kenya. Key: variable type- light blue box (mediator), rose box (neonatal mortality), white box (covariates), mint box (direct predictors of neonatal mortality), and lavender box (cluster random variable)

***Coefficients(β), all pathways are significant i.e. p<0.05**

In Tanzania, ANC attendance and SBA were mediating factors for neonatal mortality. ANC attendance mediated 2.2% of the total effects of LBW on neonatal mortality. Women who had low ANC attendance were more likely to have LBW newborns who experienced deaths. ANC attendance also mediated 5.5% of the total effects of PNC attendance on neonatal mortality. Women who had low ANC attendance were more likely not to attend PNC and have newborn deaths. The direct determinants of ANC attendance as a mediating factor were timing of ANC, quality of ANC, place of residence, maternal education and travel distance to the health facility. The results also showed ANC attendance and SBA sequentially mediated 1.0% of the total effects of PNC attendance on neonatal mortality. Women who had low ANC attendance were more likely not to have SBA, and women who did not have SBA were also more likely not to attend PNC and experience newborn deaths. The direct determinants of SBA as a mediating factor were place of residence, quality of ANC, maternal education, health insurance, and travel distance to the health facility, see Figure 7.4.

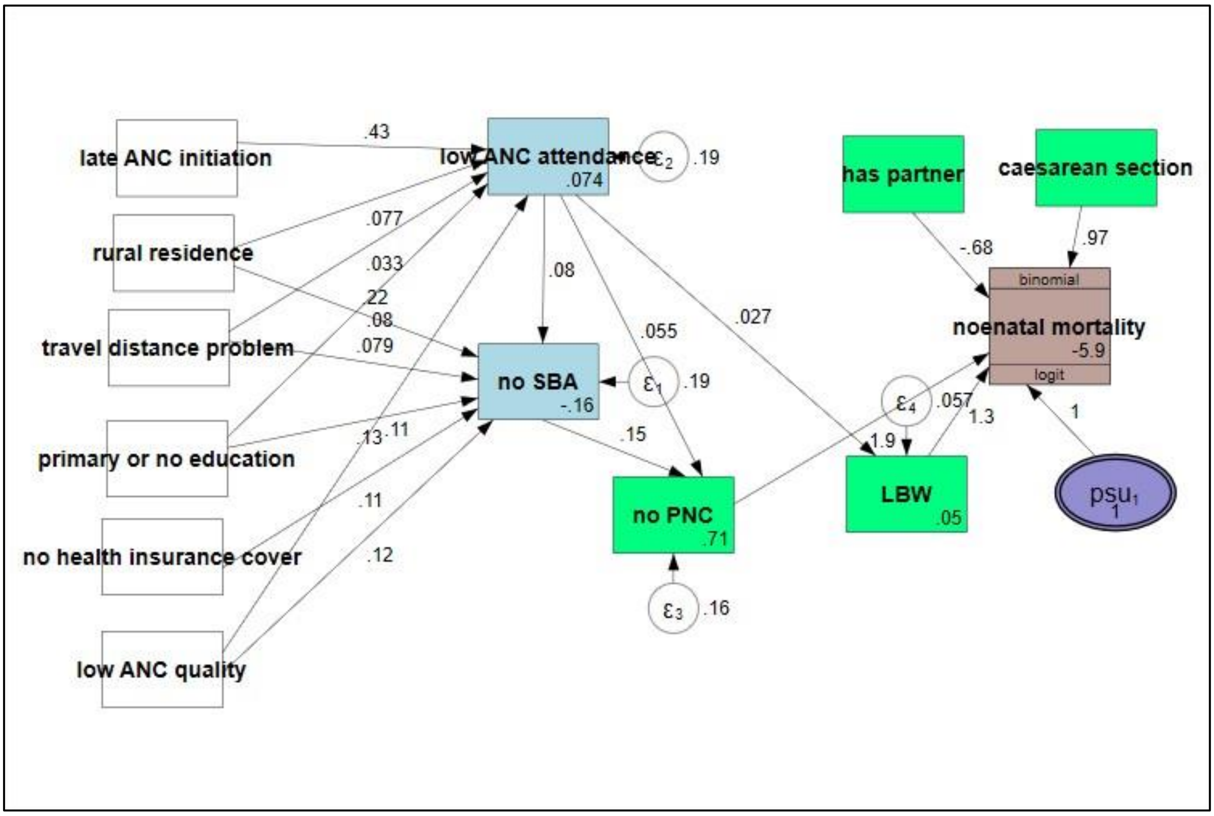

Figure 7.4: GSEM model of mediating role of MNCH services utilisation on neonatal mortality in Tanzania. Key: variable type- light blue box (mediator), rose box (neonatal mortality), white box (covariates), mint box (direct predictors of neonatal mortality) and lavender box (cluster random variable)

***Coefficients(β), all pathways are significant i.e. p<0.05**

Table 7.5: Neonatal mortality mediation analysis using the GSEM in the DRC, Kenya and Tanzania

		Neonatal Mortality			
Country	Effect of	Direct β (95% CI) (p-value)	Indirect β (95% CI) (p-value)	Total β (95% CI) (p-value)	% mediated
DRC	Parity on neonatal mortality via ANC attendance	1.00(0.43-1.57) (0.001)	0.05(0.01-0.10) (0.043)	1.05(0.45-1.65) (0.001)	4.8%
Kenya	PNC attendance on neonatal mortality via ANC attendance	1.03(0.40-1.66) (0.001)	0.10(0.03-0.18) (0.006)	1.13(0.44-1.82) (0.001)	8.8%
	PNC attendance on neonatal mortality via ANC attendance and SBA	1.03(0.40-1.66) (0.001)	0.02(0.01-0.03) (0.005)	1.05(0.41-1.69) (0.001)	1.9%
Tanzania	LBW on neonatal mortality via ANC attendance	1.31(0.53-2.10) (<0.0001)	0.03(0.01-0.06) (0.020)	1.34(0.54-2.15) (0.001)	2.2%
	PNC attendance on neonatal mortality via ANC attendance	1.90(0.80-3.00) (0.001)	0.11(0.03-0.18) (0.005)	2.01(0.84-3.17) (0.001)	5.5%
	PNC attendance on neonatal mortality via ANC attendance and SBA	1.90(0.80-3.00) (0.001)	0.02(0.01-0.04) (0.004)	1.92(0.81-3.04) (0.001)	1.0%

7.2.4 Comparison of GSEM models with or without cluster random effects

The GSEM models with random effects at cluster level had a better fit than the models without random effects based on the Bayesian information criterion (BIC) and Akaike information criterion (AIC) values. The BIC and AIC values were lower for the GSEM models with cluster random effects than GSEM models without random effects in the DRC, Kenya and Tanzania, see Table 7.6.

Table 7.6: Comparison of GSEM models with or without cluster random effects

	GSEM without random effects		GSEM with random effects at cluster level	
	AIC	BIC	AIC	BIC
Country				
DRC	23583.36	23689.79	23556.49	23670.34
Kenya	21777.77	21290.45	21742.84	21892.32
Tanzania	23640.31	23824.83	23630.41	23821.76

7.3 Discussion

The aim of the study was to elucidate the role of MNCH services utilisation in mediating neonatal mortality in three sub-Saharan African countries, namely the DRC, Kenya, and Tanzania. The study showed both direct and indirect effects of MNCH services utilisation on newborn mortality. In Kenya and Tanzania, ANC attendance mediated the total effects of PNC attendance on newborn mortality. ANC attendance and SBA were also deemed to be sequential mediators in the link between PNC attendance and newborn mortality. In Tanzania, ANC attendance also mediated the total effects of LBW on newborn mortality. In the DRC, no presence of mediation was observed, however, ANC attendance moderated the relationship between parity and neonatal mortality. GSEM findings show the interconnections between MNCH services utilisation and neonatal mortality.

Our study showed that ANC attendance mediated 8.8% in Kenya and 5.5% in Tanzania of the total effects of PNC attendance on neonatal mortality. The WHO recommends that women receive a minimum of four ANC visits during pregnancy since this period is when the fetus is affected by restricted growth, disease infections and congenital malformations which increases the chances of newborn death (3, 4). Attending ANC is an avenue for mothers to get checked on the status of their health and the fetus's health, assessed on existing and arising pregnancy complications and receive health information and advice (3). Earlier research found that ANC visits are positively linked to uptake of PNC (226). Women who had four or more ANC visits were associated with early uptake of PNC than women who had less than four ANC visits (226). This could be that that women who attended fewer ANC visits were less exposed to health education and promotion messages on managing birth complications and the significance of postpartum care. This reduces the likelihood of mother attending postnatal checks, thereby reducing the prognosis for the survival of her newborn.

Our study also found ANC attendance and SBA sequentially mediated 1.9% in Kenya and 1.0% in Tanzania of the total effects of PNC attendance on neonatal mortality. Women with less than four ANC visits had a higher probability of not having SBA, and conversely, women who did not receive SBA were more likely not to receive PNC and have higher chances of their experiencing newborn death. A potential explanation for this association could be that women who attend fewer ANC visits are less aware of complications that may arise during pregnancy and birth (227). As a result, the willingness to seek skilled care during delivery and postnatal checks after delivery is lower, which reduces the chances of their babies' survival (227). These findings also highlight the interconnections between MNCH services utilisation and neonatal mortality.

Our study also showed that ANC attendance mediated 2.1% of the total effects of LBW on neonatal mortality in Tanzania. Women who had low ANC attendance were more likely to have LBW babies, who were at a higher risk of death. Previous research has shown that neonates with LBW face a higher risk of malnutrition and childhood morbidities such as diarrhoea and pneumonia (228, 229). Frequent ANC visits are essential for monitoring maternal health and early diagnosis of danger indicators of fetal growth (3, 230). For instance, prioritising monthly measurements of body weight and fetal growth can be an effective way of monitoring the fetus' growth and lowering the risk of LBW (3, 230). Adequate monitoring of the mother's nutritional status can also reduce the risk of LBW. If a mother is iron deficient, iron tablets can be provided as a dietary supplement to both the mother and the fetus, and this lowers the risk of LBW. Thus, women who attend fewer ANC visits are more likely to miss out on these interventions, increasing the possibility of their babies being born LBW and decreasing their odds of survival.

In the DRC, our study no presence or mediation, however, ANC attendance moderated the relationship between parity and neonatal mortality. Mothers of three or more parity were more likely have low ANC attendance and have newborn deaths. mediated 4.8% of the total effects of parity on neonatal mortality. Mothers with three or more children were more likely to have less than four ANC visits and experience neonatal deaths. Previous research has shown that high parity women are less likely to attend four or more ANC visits (231, 232). This could be because multiparous women have maternal experiences of child birth and may not consider attending ANC visits frequently worthwhile (231). This negative effect on ANC utilisation, may result in more newborn deaths among multiparous women because they miss out on essential interventions provided during ANC.

The study also found that the inclusion of random effects at the cluster/community level improved the model fit in the analysis. These findings show that interventions towards reducing neonatal mortality should also be targeted in communities. Future studies should consider determining the factors associated with neonatal variation in communities to help design effective community-based interventions.

7.3.1 Strengths and Limitations

The study used nationally representative data and therefore the findings are generalisable. The study adds important empirical evidence to the body of literature on the mediation role of MNCH services utilisation on neonatal mortality. The study's limitations are that important information on maternal characteristics like gestational age and child health outcomes such as LBW, and neonatal mortality was self-reported by the mother over a long recall period and therefore subject to recall bias. However, to minimise the recall bias only information on the last pregnancy or child was utilised in the study. Additionally, child health cards were also used as sources of information if available. Another limitation is that the study did not consider risk factors of neonatal deaths by timing, and risk factors of neonatal deaths such as newborn illnesses or infections. The study is also cross-sectional and thus may not ascertain temporal causality.

7.4 Conclusions

The mediation role of MNCH on newborn mortality identifies critical areas for intervention to reduce neonatal mortality. Targeted interventions addressing barriers of low uptake of ANC, SBA and PNC should be considered to reduce newborn deaths. Furthermore, low-cost interventions such as health education and home care practices that promote the identification and management of LBW babies in low-income settings should be strengthened.

7.5 Chapter Summary

The risk of a child's death within the first month of life is high. Most causes of newborn deaths are preventable and treatable. Countries in SSA like the DRC, Kenya and Tanzania experience high rates of neonatal mortality. A crucial aspect of preventing neonatal mortality is identifying modifiable factors linked to neonatal deaths. The high burden of neonatal mortality in SSA has been linked to several modifiable factors including low ANC utilisation, low SBA and PNC uptake. However, utilisation of maternal and newborn healthcare (MNCH) services i.e. ANC, SBA and PNC remains suboptimal in SSA. The associations between MNCH services utilisation and neonatal mortality have often been determined using approaches that investigate direct associations rather than indirect associations which limits our understanding of the interconnections between different utilisation of MNCH services and neonatal outcomes. In this study, we sought therefore to breach this gap by exploring the mediation pathways of MNCH services utilisation on neonatal mortality.

The study showed both direct and indirect effects of MNCH services utilisation on newborn mortality. In Kenya and Tanzania, ANC attendance mediated the relationship between PNC attendance and newborn mortality. ANC and SBA were also deemed to be sequential mediators in the relationship between PNC attendance and newborn mortality in Kenya and Tanzania. In Tanzania, ANC attendance mediated the association between LBW and newborn mortality. Although there was no presence of mediation in the DRC, ANC attendance moderated the relationship between parity and neonatal mortality. The GSEM findings showed the interconnected associations between MNCH services utilisation and neonatal mortality. Targeted interventions addressing barriers of low uptake of ANC, SBA and PNC should be implemented to reduce neonatal mortality. Furthermore, low-cost interventions such as health education and home care practices that promote the identification and management of LBW babies in low-income settings should be strengthened.

CHAPTER 8: GENERAL DISCUSSION

This study aimed to describe the utilisation of MNCH services and associations with neonatal mortality in three sub-Saharan African countries including the DRC, Kenya and Tanzania. This chapter is a summary of consolidated findings with regards to key objectives of this thesis. The discussions of the consolidated findings of this thesis entails (1) patterns of MNCH services, (2) key emerging themes on MNCH services utilisation encompassing organisational, structural or social, and individual factors (3) linkages of MNCH services utilisation (4) the mediatory role of MNCH services utilisation on neonatal mortality. The discussions will also cover the application of statistical methods including Machine Learning and GSEM.

8.1 Summary of findings

8.1.1 Low utilisation of MNCH services

There was a general decline of women who initiated ANC late in Kenya between 2008 (67.8%) and 2014 (60.5%), and in Tanzania between 2010 (60.9%) and 2016 (49.8%), while the DRC showed contradicting temporal patterns between 2007 (56.8%) and 2014 (61.0%)). Although, the trends in Kenya and Tanzania indicated a decline in late ANC uptake, uptake of early ANC is relatively low. Very low levels of the utilisation of CoC in terms of attendance four or more ANC visits, SBA and PNC in the DRC, Kenya and Tanzania were also reported in this study. Over 70% of the women and children in the DRC (91.0%), Kenya (72.4%) and Tanzania (93.6) were not retained or dropped out of the MNCH continuum. Most women and children dropped out from CoC during the postpartum period, which placed them at greater risk of experiencing adverse health outcomes.

The consolidated findings in line with the objectives of the study are outlined in Table 8.1.

Table 8.1: Consolidated Study Findings

Chapter	Objectives	Key findings
4	To assess the trends and associated determinants of late ANC initiation in three in three sub-Saharan African countries: the DRC, Kenya, and Tanzania, from 2007 to 2016.	The study results showed declining trends in late ANC initiation over time in Kenya and Tanzania, except for the DRC. Inequalities in late ANC uptake were observed among various geographic, educational, parity, and economic groups.
5	To predict the main risk factors of PNC non-utilisation in three sub-Saharan African countries using the Decision Tree.	The Decision Tree models exhibited higher accuracy in predicting non-utilisation of PNC than the Logistic Regression models. Using the Decision Tree, women who had low quality ANC, home deliveries and unemployment had the highest probability of not utilising PNC (92.0%) in the DRC. In Kenya, women who were unemployed, delivered at home and had no exposure to mass media had the highest likelihood of not utilising PNC (87.0%). In Tanzania, women who had home deliveries, low quality ANC and unwanted pregnancies showed the highest likelihood of not utilising PNC (100.0%).
6	To predict the likelihood of a mother/child dropping out from the MNCH continuum and determining the most influential predictors in three sub-Saharan African countries.	The study found high rates of dropping out from the MNCH continuum in the DRC, Kenya and Tanzania. The largest drop out from the CoC occurred during the postpartum period. Common predictors of the drop out from the MNCH continuum were rural residence, lower maternal education, and non-exposure to mass media. Further, the influence of factors such as maternal age, parity, access to money for medication, travel distance to the health facility, household wealth, household size varied by country. The Random Forest had superior predictive accuracy compared to other Machine Learning models. Household wealth, place of residence, maternal education and exposure to mass media were ranked among the top four predictors of the drop out from the MNCH continuum.
7	To describe the mediation role of MNCH services utilisation on neonatal mortality in three sub-Saharan African countries.	ANC attendance and SBA were found to be mediators of neonatal mortality. ANC attendance mediated the total effects of PNC attendance on newborn mortality by 8.8% in Kenya and 5.5% in Tanzania. ANC attendance and SBA sequentially mediated the total effects of PNC attendance on newborn mortality by 1.9% in Kenya and 1.0% in Tanzania. In Tanzania, the findings also showed that ANC attendance mediated 2.8% of the total effects of LBW on newborn mortality. No presence of mediation was observed in the DRC, however, ANC

		attendance moderated the relationship between parity and neonatal mortality.

8.1.2 Emerging themes on MNCH services utilisation

The key merging themes on the utilisation of MNCH services utilisation encompassed structural or social, organisational or health system and individual factors.

8.1.2.1 Structural or social factors

The main structural or social factors linked with the utilisation of MNCH services in the study were household wealth (poverty) and place of residence.

Women from low-income households were more likely to start ANC late in the DRC, Kenya, and Tanzania. Inequalities in late ANC uptake also continue to increase between high-and-low income households in Kenya and Tanzania. Low-income households were also associated with increased odds of dropping out of the MNCH continuum in Kenya and Tanzania. Women of low socio-economic status struggle to access care due to high medical consultation fees, costs of buying medicines and transport costs (212, 233). The low socioeconomic status of women also affects their decision making in seeking maternal healthcare (234). In some regions, men oversee most of the household finances, making it difficult for women to get medical care or transportation to healthcare facilities and limiting women's ability to make health-related decisions (234). Although Kenya and Tanzania have polices on free maternal health (87, 153), many women in impoverished circumstances pay OOP fees to cover medical costs (87, 153).

Rural women were also more likely to commence ANC late in the DRC and Kenya. Rural-urban inequalities in late ANC uptake also continue to widen in the DRC and Kenya. Rural women were also associated with increased odds of dropping out from the MNCH continuum in the DRC, Kenya and Tanzania. Rural women are mostly likely to face more barriers to accessing maternal healthcare than their urban counterparts resulting in low utilisation of maternal healthcare. Cultural factors, low socioeconomic status and limited education are obstacles to rural women's uptake of MNCH and utilisation of the CoC (163, 234, 235). Elsewhere cultural norms affect women's autonomy in making decisions. One of the cultural problems that contributes to late ANC uptake or no uptake at all is the tradition or belief in concealing pregnancy (236). Some African women believe that showing pregnancy at an early stage could harm the fetus as the baby is vulnerable and at risk of being subject to "witchcraft (236). They would hide their pregnancy in the early months to avoid being bewitched and thus protect their babies. Such cultural beliefs result in delays in seeking ANC (236). Secondly, women in some African cultures rely upon their mothers-in-law or elderly peoples' decisions with regards to seeking care, especially those who marry early, and can be given inappropriate counsel on when and where to seek care (237) .

Community peer influence is another challenging cultural aspect. Most African families in the rural areas are in a setup of a community and there a exist community/peer influence from the community members' sharing their opinions (negative one) about the state of maternal health with other women (238). Such conversations happen in community gatherings or interactions. For instance, some women may share their traditional birth attendance stories and convince other pregnant women to consider using traditional birth attendants and neglect using skilled birth attendants (239).

The low socioeconomic status of rural women creates strong financial dependency on their male partners throughout their lives which hinders them from deciding about their own health and wellbeing. In some African settings, women need to seek approval to visit health facilities and may not be permitted to go alone (240). Lack of financial means to reach to health facilities due to transport costs and pay for services, and the preoccupation of supporting a large household on a limited finances also hinders rural women of low socioeconomic status from utilising MNCH (241).

Low levels of education among rural women have a significant effect on the level of their decision-making power at household level. Rural women with low levels of education attainment have low confidence in decision making (242) and could be less capable of obtaining a gainful employment, which affects their economic contributions to the household and decision making (242).

A multifaceted approach is required to overcome obstacles to receiving timely and adequate utilisation of MNCH among women living in poverty settings and rural areas. Interventions that mitigate cultural norms, give women autonomy in decision making, reverses traditional practices, promote girls and women economically and improve access to healthcare are crucial in MNCH. Women's autonomy must be supported, and women should be empowered to make their own health decisions through dedicated health education and promotion programs that target pregnant women, their families, and the entire community (163). Some of these interventions include rural educational campaigns which provide educational accessibility to empower women in their health rights, benefits and advantages of utilising health services over traditional approaches. Such knowledge will empower the women to be independent and regain their autonomy when it comes to their health.

In addition, community-based interventions such as income generating projects like women clubs, that empower low socioeconomic rural women financially are commendable. These income generating projects will provide economic stability improve the women's engagements and decision making. Social marketing, media efforts, use of religious/cultural leaders' interventions and formal education for girls can also help mitigate some of the cultural norms experienced by rural women (234).

The use of Community Health Workers (CHWs) can facilitate healthcare access by using CHWs as a liaison between healthcare providers and rural residents to ensure their healthcare needs are met (243). Freestanding Emergency Departments defined by the American College of Emergency Physicians as a “facility that is structurally separate and distinct from a hospital and provides emergency care can be used to provide a range of different speciality care for rural communities that lack hospital inpatient care (243).

Provision of an adequate workforce is necessary for maintaining healthcare access in rural communities. In order to increase access to healthcare, rural communities must use their healthcare professionals in the most efficient and strategic ways (243). This might include allowing each professional to work at the top of their license, using new types of providers, working in interprofessional teams, and creative scheduling to offer clinic time outside of regular work hours (243).

8.1.2.2 Organisational or health system factors

The main organisational or health system factors linked with the utilisation of MNCH services in the study were travel distance to the health facility, medical costs, place of delivery and quality of ANC.

The study found that women who faced big problems of travel distance to the health facilities were linked to higher odds of dropping out of the MNCH continuum in Kenya. The challenges of long travel distances to health facilities are common in SSA, particularly in rural areas (176, 244, 245). The low geographic accessibility of health distance is linked to poor roads and inadequate health facilities (176, 244, 245). Moreover, nearby health facilities in rural areas are extremely understaffed with limited medical equipment and supplies and well-equipped health facilities and thus women are forced to consider facilities which are further away (246). Even relying on public transport alone is not sufficient as there might be delays. At times, the public transport is unaffordable and not available due to poor roads and rough terrains (176).

The study also revealed that women who faced challenges with medical costs were linked to the high drop out of the MNCH continuum in Kenya. Service-related costs deter the utilisation of ANC, institutionary deliveries, and PNC at health facilities (247, 248). Unaffordable service-related costs are still a problem even in countries were policies guaranteed free maternity services like Kenya (88). Many women in Kenya still incur OOP payments such as registration fees, hospital card, consultation, admission, lab tests, ultrasound, payment for surgery (caesarean section or any theatre fees), blood and drugs, and other related costs (88). Other OOPs expenses include cost of payment for items that the patient is told to buy like cotton wool, basin, bandages/gauze rolls, syringe, drugs and costs such as, accommodation cost for self and a companion while the patient is admitted, and any form of informal payments. The high amounts of OOP payments reflect Kenya's health financing system, which is mainly based on user fees at the time-of-service delivery (246).

The study also showed that that women who gave birth at home were more likely not to receive PNC after delivery in the DRC, Kenya and Tanzania. Women who give birth at home often receive care from traditional birth attendants (TBAs) (169). TBAs are regarded as well-known, trustworthy, and respected custodians of traditional knowledge who actively engage with women during pregnancy, birth, and, especially, postpartum (169, 249). Other women also prefer TBAs to supervise their childbirth due to previous bad interactions with healthcare professionals (HCPs). In their view TBAs are more sensitive, caring, hospitable, affectionate, and carried a more positive presence than healthcare professionals (HCPs) (250). However, TBAs lack the necessary expertise, skills, and equipment to identify and address obstetric and neonatal problems. When issues occur at home, organising referral and transport to the health institution is a challenging process for everybody involved, and likely to result in significant delays or not seeking PNC, particularly in rural areas (249). Culture also plays a major role in the way a woman perceives and prepares for her birthing experience (251). The cultural practice of keeping newborns indoors results in negligence of seeking PNC (251).

The term "quality of care" is a holistic expression that implies a host of features such as the range of medical services offered, clinical competence of the staff, hospital amenities, expertise of the physicians, ambience of the hospital, staff behavior, most importantly, patient satisfaction. Quality of care is likely to improve pregnancy and childbirth outcomes if uptake of services is high (183, 184). Our findings showed that women who received low quality of ANC in terms of routine examinations were more likely to lack PNC in the DRC, Kenya and Tanzania. Women are discouraged from accessing care in health institutions due to the low quality of care provided (58). In some situations, HCPs are not available to assist them when they book appointments, or, more frequently, there is no medical equipment accessible to provide them with the essential care. As a result, they are less inclined to seek care at health facilities (58). At times, the attitudes of HCPs are rude, abusive, insensitive, or deliberately negligent which may result in women seeking care at home (250).

The number of specialists and the duration of their presence is limited in rural health centres. In some settings, women do not attend care at health institutions or return before being attended due to long waiting times. Long waiting makes women feeling bored, tired, or as if they were wasting their time. Long waiting times are affected by staff shortages, lack of equipment, lack of examination space and lack of prioritisation of PNC. Privacy is another factor that explains women's preferences of utilising healthcare facilities (250). Facility-based services that provided privacy are preferred by many women who were concerned about giving birth in open settings and having private parts exposed to strangers (250). Ensuring privacy can also promote more effective communication between the healthcare worker and patient, which is essential for quality of care (252).

Poor referral systems, particularly in rural areas, have an impact on health-care delivery (253). The number and duration of specialists' presence in rural health centers are limited. Overcrowding often compels patients to wait a lengthy time or visit a specialist's private office. The referral system is intended to optimise the usage of three levels of health care while avoiding excessive congestion and waste of people and material resources in the specialized levels. However, limits in specialist levels and an overburden of non-urgent recommended cases have rendered the referral system ineffective (253).

The choice of utilising care at health facilities is an important decision that is affected by various organisational, or health system factors related to geographic accessibility, affordability, and quality of care. Interventions aimed at addressing all facets of organisational or health-care system barriers to accessing MNCH should be prioritised if the SDG goals are to be met. Governments need to create strategies that make health facilities more accessible, especially for women living in remote (rural) areas, these include increasing health facilities, setting mobile clinics, providing alternative transportation such as motorcycle ambulances and improving the roads infrastructure (176, 244). Government funding, such as revenue collection, risk pooling, and purchasing, are essential strategies in eliminate direct OOP payments and reducing medicals costs (254). Policymakers should consider integrating TBAs into the formal health system to eliminate the financial, geographical, and cultural barriers to care-seeking outside the home during the postnatal period (255). Policymakers, managers, and healthcare practitioners should enhance the quality of care offered by allocating more resources such as staff, equipment and medical supplies to health facilities (255).

8.1.2.3 Individual factors

The main individual factors linked with the utilisation of MNCH services in the study were maternal education, exposure to mass media, maternal age, parity, employment status and pregnancy intention.

Illiterate or low educated women were more likely to have late ANC initiation in the DRC, Kenya, and Tanzania. Women with low education were also more likely to drop out from the MNCH continuum in the DRC, Kenya and Tanzania. The lack of adequate knowledge regarding the use of MNCH is a significant and plausible factor for its low utilisation among illiterate or low educated women (256). Women with low levels of education are also less empowered in decision-making because they lack the confidence to exercise autonomy and they are more likely to be financially dependent, limiting their access and use of healthcare services (257).

Limited exposure to mass media was also associated with a high drop out from the MNCH continuum in the DRC, Kenya and Tanzania. Newspapers, radio, and television are media outlets through which important health information is transmitted to women. However, in rural areas in most developing countries have limited connectivity and access to televisions, radios and online health information (258). In this sense, women who do not have access to media do not have easy access to information that would assist them make informed decisions about their health and improve their tendency to use MNCH services (259). Exposure to mass media plays an important role in creating awareness on reaching vulnerable populations of less educated and rural mothers (259).

High parity women were more likely to initiate ANC late in the DRC, Kenya and Tanzania. High parity was also linked with dropping out from the MNCH continuum in the DRC and Kenya. This may be due to the fact that women who have higher parity have more experience in pregnancy and childbirth, they become more confident to stay at home and less interested to take up MNCH services and utilise the CoC (28, 260). The study also showed that older women aged 25-49 years in the DRC and Kenya had less odds of dropping out of the MNCH continuum. This may be because older women prioritise MNCH more than younger because of accumulation of knowledge which comes with age or birth-related complications and poor health conditions which are more common as age advances which trigger the women to go to health facilities (261).

Unemployed women had higher odds of not attending PNC in Kenya. Women have less job opportunities than men in Kenya. Long-term unemployment among women in Kenya is higher (63.6%) for women without formal education, compared to men (18.7%) (175). Unemployment is also higher among women with an academic background beyond secondary school (25.8%) compared to men (11.7%) with the same level of education (175). A woman's unemployment is viewed as a barrier to healthcare utilisation. Financial issues have been linked to the difficulty of women to pay for medications, transport costs, vaccines, and healthcare for their infants (171, 262). Women who are financially dependent are also more likely to have less autonomy in making decisions with their male partners, even decisions which pertain their health (32).

Unintended pregnancies are pregnancies that are mistimed, unplanned or unwanted at the time of conception. The study demonstrated that women who no longer wanted to be pregnant were associated with lack of PNC in Tanzania. A plausible reason could be that a mistimed or an unwanted pregnancy is most often identified later than a wanted pregnancy and could result in women having less time to access ANC services, affecting the utilisation of later services such as PNC (185, 186). Other plausible reasons could be that women experiencing an unintended pregnancy are more likely to be socio-economically disadvantaged, have no formal education, and have lower autonomy which limits their capability to utilise recommended MNCH services because of either financial difficulties or poor knowledge regarding the importance of MNCH use.

Interventions designed at the individual level to improve uptake of MNCH should target various population groups. Government policies supporting female education can enhance women's understanding of the significance of utilisation of MNCH services. MNCH educational programs which are delivered via accessible mass media channels such as short audio and video clips using local languages can improve the utilisation of MNCH services (178). Educational programs on MNCH that target high-parity women and various age groups (young and old) and promote community knowledge about the importance of MNCH should be prioritised. Governments should consider policies that create more job opportunities for women to equip them economically. Health programmers and stakeholders should also consider integrating MNCH services with family planning services for early detection of women having an unintended pregnancy (263).

8.1.3 Linkages of MNCH services utilisation

The GSEM results showed the interconnections between timing of ANC, ANC visits, SBA and PNC. Women who started ANC late were more likely to attend less than four ANC visits in the DRC, Kenya and Tanzania. Women who attended less than four ANC visits were also more likely to have SBA and PNC in Kenya and Tanzania. These findings underline the necessity of enrolling women in ANC as early as possible, as this significantly affects their use of services in the CoC (264). An early single contact with the health facility during pregnancy influences the use of MNCH services throughout the pregnancy. Perhaps we could explain these findings by pointing out that women who attend ANC early are more likely to have more ANC sessions. During ANC sessions, women receive a series of health education, and different cadres of healthcare providers sensitise and discuss with women on the necessity of skilled care during delivery and postpartum period to prevent pregnancy and birth related complications (265).

8.1.4 Comparison of the Machine Learning and Logistic regression

The logistic regression is one of the most used statistical methods (266). It is simply and easier to interpret than complex Machine Learning models such as Artificial Neural Networks, Support Vector Machines, and Random Forests (266). Although Machine Learning algorithms are used in several research fields, their complicated structure makes them difficult to interpret.

The output of logistic regression is usually in the form of probabilities, which makes it easier to understand than other black-box Machine Learning models that output complex relationships between input and output variables (266). An interpretable model assists the human understanding of the prediction process and avoids false assumptions which improves the decision-making process (266). The logistic regression is also easier to train or execute than complex Machine learning algorithms (266) because it requires few computation resources and small data as compared to complex Machine Learning algorithms. It also has fewer parameters to tune compared to other Machine Learning algorithms. The logistic regression also focuses on understanding causality in comparison to complex Machine Learning algorithms which focus on prediction and classification (266, 267).

Simpler models like logistic regression offer easier interpretability, however, complex Machine Learning algorithms can achieve high levels of accuracy (267). The accuracy of the Machine Learning algorithms can be improved by training the model using different parameters. However, the performance of complex Machine Learning algorithms can be affected by class imbalance of the outcome variable (115, 267). Compared to the logistic regression, complex Machine Learning algorithms are also better in handling larger data and a larger number of potential predictors and uncovering hidden patterns (266). However, Machine Learning algorithms lack inbuilt functions to account for clustering and stratification for complex survey data such as DHS surveys (268) .

The findings in our study support previous research (78). The Decision Tree model had higher performance compared to the Logistic Regression model in predicting the risk factors of non-utilisation of PNC. The main advantage of the Decision Tree consists of its ability to describe patterns in the data by revealing high order interactions among variables, whereas the Logistic Regression just shows important interactions, without ordering them (78). The data from the Decision Tree results is turned into a tree of decision rules that can be easily interpreted and understood. The data from the Decision Tree model can be displayed in form of simple and visual tree graphics that can be easily applied in public health practice (78). Application of high accuracy Machine Learning approaches with easier interpretability, such as the Decision Trees, can aid in identifying the main factors contributing to the low utilisation of PNC, as well as other MNCH services in SSA. This approach is useful in prioritising interventions in high-risk sub-populations, thereby optimising public health resources and improving the uptake of MNCH in SSA.

The study also applied several complex Machine Learning models such as the Random Forest, Artificial Neural Network and Support Vector Machine to predict the drop out from the MNCH continuum. Application of several Machine Learning models offers new opportunities for improving existing methods for predicting MNCH services utilisation (12). Conventional statistical models such as the logistic regression have traditionally been used to examine data and identify risk factors, however these methods are limited in reflecting complicated connections. Machine Learning can account for interactions between multiple components, handle non-linear correlations, and adapt to changing patterns in data (12). For instance, the Random Forest allows an unlimited number of variables to be incorporated into the model. The algorithm automatically tests several hypotheses and selects features that best predicts the outcome using information gained from each variable (79).

The study findings showed that Machine Learning classification models such as the Random Forest, Decision Tree, Artificial Neural Network and Support Vector Machine had higher predictive accuracy of the drop out from the MNCH continuum compared to the Logistic Regression. The study shows the superior capabilities of Machine Learning methods compared to other conventional approaches in predicting MNCH services utilisation. Other similar research in MNCH have been shown that Machine Learning methods outperform conventional statistical methods (34).

The study findings also showed that the Random Forest had the highest prediction accuracy, followed by the Decision Tree. The results showed that the Random Forest was best suited for predicting the discontinuation of MNCH. When compared to other Machine Learning algorithms such as Artificial Neural Network and Support Vector Machine, the Random Forest is more interpretable and adaptable (35). However, it is essential to note that the performance of other Machine Learning models may differ depending on the dataset and the research problem being addressed. Therefore, it is recommended to assess the performance of multiple models and select the one that performs the best for a given problem (12).

Using the Random Forest algorithm, the study ranked the most important predictors of the drop out from the MNCH continuum. The use of Machine Learning to identify and rank the most influential predictors can enhance improvements in the utilisation of the MNCH continuum in the SDG era. Public health programmers and policymakers can utilise this information to design and implement cost-effective interventions which are timely and optimise use of public health resources (216). Machine Learning can also be used in the development of rapid response mechanisms such as web-based programs which can predict whether a pregnant woman and her unborn child will be retained in the CoC depending on the maternal characteristics (34, 217). This allows for real-time tailored interventions for pregnant women at high risk of dropping out of care and thus promoting retention in the MNCH continuum (34, 217).

8.1.5 Application of the GSEM

GSEM describes causal models in an integrated approach. It is a system for specifying the interrelationships among observed and latent variables with greater specificity than simpler methods such as multiple regression. There are so many observed and latent variables inter-linked in MNCH research and the GSEM can be used to test hypotheses of different MNCH services and health outcomes and their predictors based on the theoretical model of the study (131). The GSEM established interconnections between MNCH services utilisation such as timing of ANC, ANC visits, SBA, PNC and neonatal mortality. This suggests that more than direct factors are accountable for the associations between MNCH services utilisation and neonatal mortality. This knowledge is essential to develop policy alternatives for tackling the challenges that SSA faces in its efforts to enhance newborn survival through timely and appropriate use of MNCH services.

8.2 Recommendations for Future research projects

Due to the cross-sectional nature of our data, we could not establish temporal causality. Therefore, we recommend longitudinal studies which assess the utilisation of MNCH services of women at different periods from pregnancy to postpartum period, and their health outcomes. Most of the information in our study was self-reported and subject to recall bias. Therefore, we recommend future studies that utilise sources of information that minimise recall bias such as patient records. In addition, we recommend more studies that investigate how health facility factors such as health worker to patient ratio, health facility operations, availability of drugs and medicines affect the utilisation of MNCH services. There is minimal research on the impact of health facility service provision on MNCH services utilisation, particularly in SSA, owing to a lack of health facility data (269). Studies which utilise health facilities data linked individual/patient records are useful for describing the quality-of-service provision by health facilities in SSA and revealing gaps in service provision across health facilities, as well as its influence on MNCH services utilisation.

We have several recommendations regarding the implementation Machine Learning. Firstly, the application of Machine Learning is still growing, particularly in SSA, therefore there is a need for capacity building through training and provision of technological resources to increase its use. We also suggest collaboration between research institutions, academic stakeholders, policy makers and regulatory authorities for exchange of knowledge and information. Secondly, the capabilities of Machine Learning application in research are still developing. For instance, our study could not consider survey design in the Machine Learning analysis because of the lack of inbuilt functions in Machine learning algorithms to incorporate clusters and stratification variables for complex survey data such as DHS surveys and this may limit generalizability of the findings (268). Hence, we recommend more research of growing the capability functions of Machine Learning methods in all Machine Learning softwares.

8.3 Strengths and Limitations

The strength of this study is that the study used nationally representative data therefore the findings are generalisable and applicable to countrywide policies and interventions. In addition, the study explored the use of novel analytical methods like Machine Learning in predicting the utilisation MNCH services. These methods are computationally strong than conventional statistical methods when using big data like DHS surveys and capable of revealing hidden patterns and relationships in data. The study also adds empirical evidence on the interrelationships between utilisation of different MNCH services and neonatal mortality to the body of literature.

However, the study has several limitations. Questions on main services of MNCH, gestational age, childbirth weight and newborn mortality relied on women's self-report which is subject to recall bias. Our research also depended on the completeness of data on variables of interest in all rounds of the DHS surveys; as a result incomplete or unavailable data, the study did not evaluate other organisational factors such as quality of obstetric and postnatal care, structural factors such as traditional practices and religion, relationship factors such as family support and relationship power dynamics, and individual factors such as history of pregnancy and obstetric complications. Additionally, the generalisability of the Machine Learning analysis results is limited because the analysis could not factor the complex survey design (268). Additionally, it was also not possible to establish temporal causality due to the cross-sectional nature of the data.

8.4 Conclusions

In conclusion, this study aimed to describe the utilisation of the MNCH services and associations with neonatal mortality in three sub-Saharan African countries namely the DRC, Kenya and Tanzania. Overall, the study showed suboptimal uptake of early ANC, very low retention of women and children in the MNCH continuum, with most women and children dropping out from the CoC at postpartum. The findings also showed the existence of social, health system and individual inequalities in MNCH and their impact on child health outcomes. Poverty, rural location, long travel distances to access health facilities, unaffordable medical expenses, home deliveries, poor quality of care, low levels of education, high parity, younger age, unemployment, limited exposure to mass media, and unplanned pregnancies characterise women who are vulnerable to unequal and poor MNCH services utilisation. Context-specific intervention programs such as female education and economic empowerment, MNCH promotion programs via mass media, improving geographic accessibility and quality of care, particularly for the most vulnerable groups of the populations such as women of low socioeconomic status and rural women are essential to improve the utilisation of MNCH services and meeting the SDG-3 goals.

Modern biostatistical methods were used to model the utilisation of MNHC services. Machine Learning methods provide essential tools to understand public health problems and should be applied in assessing the utilisation of MNCH services to develop effective targeted interventions in high-risk populations. The GSEM established interconnections between the timing of ANC, ANC visits, SBA and PNC and neonatal mortality. The timing of the first ANC contact is an important entry point to a continuation through the COC. It makes women better informed about pregnancy and the subsequent use of MNCH services. All stakeholders should direct more efforts towards promoting early ANC initiation through various initiatives including increasing women's access to information platforms, enhancing female education, improving household incomes, and bringing services closer to communities.

APPENDIX
STATA DO files

Descriptive and Multilevel logistic regression analysis

```
use drc_descr, clear

***survey set the data, using cluster, sample weight and stratification variables

svyset psu [pweight=sampwt], strata(strata)

***descriptive  analysis

svy: tab residence

svy: tab residence, count

svy: tab age_group_new_cat

svy: tab age_group_new_cat, count

svy: tab education_1

svy: tab education_1, count

svy: tab birth_order

svy: tab birth_order , count

svy: tab household_size

svy: tab household_size , count

svy: tab wealth

svy: tab wealth , count

svy: tab anc_timing

svy: tab anc_timing , count

***multilevel logistic regression model

svy:logit  anc_timing i.residence i.education_1 i.birth_order i.wealth i.household_size
i.age_group_new_cat, or

*****prediction scores

svy:logit  anc_timing i.residence i.education_1 i.birth_order i.wealth i.household_size
i.age_group_new_cat

predict p
```

GSEM mediation analysis

**GSEM model with random effects

```
gsem (rh_sba -> rh_pnc, ) (rh_anc_visits -> rh_sba, ) (rh_anc_visits -> rh_pnc, ) (rh_pnc -> neomort, family(binomial) link(logit)) (anc_timing -> rh_anc_visits, ) (residence -> rh_sba, ) (M1[psu] -> neomort, family(binomial) link(logit)) (medgo -> neomort, family(binomial) link(logit)) (medcost -> rh_pnc, ) (mother_education -> rh_anc_visits, ) [pweight = sample_weight], covstruct(_lexogenous, diagonal) latent(M1 ) nocapslatent
```

***assessing model fit

```
estat ic
```

****estimation of direct total, indirect and direct effects

```
gsem, coeflegend
```

***direct effect of PNC on neonatal mortality

```
nlcom _b[neomort:rh_pnc]
```

***indirect effects via ANC visits and SBA

```
nlcom  _b[rh_sba:rh_anc_visits]*_b[rh_pnc:rh_sba]* _b[neomort:rh_pnc]
```

**total effects

```
nlcom _b[neomort:rh_pnc] + _b[rh_sba:rh_anc_visits]*_b[rh_pnc:rh_sba]* _b[neomort:rh_pnc]
```

***indirect effects via ANC visits

```
nlcom  _b[rh_pnc:rh_anc_visits]* _b[neomort:rh_pnc]
```

**total effects

```
nlcom _b[neomort:rh_pnc] + _b[rh_pnc:rh_anc_visits]* _b[neomort:rh_pnc]
```

Machine Learning in R

```
###call packages from library
library(haven)
library(tidyverse)
library(magrittr)
library(CHAID)
library(partykit)
library(caret)
library(dplyr)
library(devtools)
require(devtools)
library(DMwR)
library(ROSE)
library(rpart)
library (rpart.plot)
library(RColorBrewer)
library(rattle)
library(vip)
library(MLeval)
library(pROC)
library(LogicReg)
devtools::install_cran("survey")
library(survey)

####bring in data
drc_mac_data <- read_dta("drc_mac.dta")
str(drc_mac_data)
head(drc_mac_data)



#Convert to factor level
drc_mac_data <-drc_mac_data %>% mutate(householdwealth=factor(householdwealth, levels = c(1, 2, 3),
labels = c('middle', 'poor', 'rich')),
      birthorder=factor(birthorder, levels = c(1,2,3), labels = c('threeormore', 'one', 'two')),
      ancvisits=factor(ancvisits, levels = c(0,1), labels = c('lessthanfour', 'fourormore')),
      SBA=factor(SBA, levels = c(0,1), labels = c('hadSBA', 'noSBA')),
      PNC= factor(PNC, levels = c(0,1), labels = c('hadPNC', 'noPNC')),
      placedelivery= factor(placedelivery, levels = c(0,1), labels = c('facility', 'home')),
      anctiming = factor(anctiming, levels = c(0, 1), labels = c('early', 'late')),
      qualityanc = factor(qualityanc, levels = c(0, 1), labels = c('low', 'high')),
      education = factor(education, levels = c(0, 1), labels = c('high', 'low')),
      househead = factor(househead, levels = c(0, 1), labels = c('female', 'male')),
      maternalage = factor(maternalage, levels = c(0, 1), labels = c('15to24years', '25to49years')),
      media  = factor(media, levels = c(0, 1), labels = c('noexposure', 'exposed')),
      health_insur = factor(health_insur, levels = c(0, 1), labels = c('notcovered', 'covered')),
      medcost = factor(medcost, levels = c(0, 1), labels = c('nobigproblem', 'bigproblem')),
      meddist= factor(meddist, levels = c(0, 1), labels = c('nobigproblem', 'bigproblem')),
      medgo = factor(medgo, levels = c(0, 1), labels = c('nobigproblem', 'bigproblem')),
      wantedpreg = factor(wantedpreg, levels = c(1,2,3), labels = c('then', 'late', 'nomore')),
      financialdecisions = factor(financialdecisions, levels = c(0,1), labels = c('hasautonomy', 'noautonomy')),
       residence = factor(residence, levels = c(0, 1), labels = c('urban', 'rural')),
      employed = factor(employed, levels = c(0, 1), labels = c('no', 'yes')),
      haspartner=factor(haspartner, levels=c(0,1), labels=c('no', 'yes')),
      caesar=factor(caesar, levels=c(0,1), labels=c('no', 'yes')),
      termpreg = factor(termpreg, levels = c(0, 1), labels = c('nohistory', 'hashistory'))) %>%
       na.omit()
```

```
#####keep the selected variables
drc_mac_data <- drc_mac_data %>%
  select(c(PNC, anctiming,placedelivery, ancvisits, qualityanc, education,
          residence,medcost,health_insur,caesar, employed,financialdecisions,weight))

str(drc_mac_data)

####split into training and testing data
set.seed(1)
# Step 1: Get row numbers for the training data
trainRowNumbers <- createDataPartition(drc_mac_data$PNC, p=0.80, list=FALSE)

# Step 2: Create the training dataset
drctrainData <- drc_mac_data[trainRowNumbers,]

# Step 3: Create the test dataset
drctestData <- drc_mac_data[-trainRowNumbers,]

#####Use ROSE correct for class imbalance
rose <- ROSE(PNC~., data = drctrainData, N = 6990, seed=111)$data
table(rose$PNC)



####Train the Decision Tree

set.seed(12)
fit.tr2 <- rpart(PNC~., data = rose, method = 'class',
              control=rpart.control(minsplit =3,
                                    minbucket =3, cp=-1, maxdepth=3))

##plot the Decision Tree
par(xpd = TRUE)

rpart.plot(fit.tr2, extra = 106)

rpart.plot(fit.tr2)

print(fit.tr2)

fancyRpartPlot(fit.tr2, caption = NULL)
text(fit.tr2, use.n = TRUE, all = TRUE)

print(fit.tr2)
###test the decision tree using test data

p <- predict(fit.tr2, drctestData, type = 'class')

confusionMatrix(p, drctestData$PNC, positive='noPNC')

predict.p<-predict(fit.tr2,newdata=drctestData, type="raw")
confusionMatrix(p, drctestData$PNC,
              positive="noPNC", mode="everything")

library(pROC)
predicttree <- predict(fit.tr2, drctestData, type = 'prob')
pROC::roc(drctestData$PNC,predicttree[,2])

roc_object <- roc( drctestData$PNC, predicttree)
```

```
###Train the logistic regression
set.seed(13)

mylogit <- train(
  PNC ~ .,
  data = rose,
  method = "glm",
  family = "binomial",
  trControl = trainControl(method = "cv", number = 10)
)

pred_class <- predict(mylogit, drctestData)



confusionMatrix(
  data = relevel(pred_class, ref = "noPNC"),
  reference = relevel(drctestData$PNC, ref = "noPNC")
)


predictlog <- predict(mylogit, drctestData, type = 'prob')
pROC::roc(drctestData$PNC,predictlog[,2])

roc_object <- roc( drctestData$PNC, predictlog)
```

Machine Learning in Python

```
###Call in packages
import pandas as pd
import numpy as np

import os

from pandas import read_stata
from numpy import set_printoptions
from sklearn.feature_selection import SelectKBest
from sklearn.feature_selection import f_classif

from imblearn.over_sampling import KMeansSMOTE
from sklearn.cluster import KMeans

from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import roc_auc_score
from sklearn.metrics import precision_score, recall_score, f1_score, accuracy_score

from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
from sklearn.model_selection import train_test_split

from sklearn.preprocessing import StandardScaler

from sklearn.preprocessing import MinMaxScaler

from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay

from sklearn.svm import SVC

from sklearn.model_selection import GridSearchCV
from sklearn.neural_network import MLPClassifier

import random
from sklearn.tree import DecisionTreeClassifier

from sklearn.ensemble import RandomForestClassifier

from sklearn.model_selection import RandomizedSearchCV

from sklearn.metrics import roc_curve, auc

import matplotlib.pyplot as plt
import numpy as np
%matplotlib inline
from sklearn.metrics import roc_curve
from sklearn.metrics import roc_auc_score
from sklearn import metrics

from sklearn.inspection import permutation_importance
from matplotlib import pyplot


print(os.getcwd())

C:\Users\chena
print(os.listdir())
```

###read the data

drc = pd.read_stata('drc_mac_data.dta',columns=['residence',
'education','birthorder','agegroup','exposuremedia',
'completeCOC'])

***create dummy variables of selected features

drc2=pd.get_dummies(drc, columns =['residence', 'education','birthorder','exposuremedia', 'agegroup',
])

***assign feature and target columns

feature_cols=list(drc2.columns[1:12])

target_cols=drc2.columns[0]

print("Feature columns:n\{}", format(feature_cols))

Feature columns:n\{} ['residence_urban', 'residence_rural', 'education_lower', 'education_higher', 'birthorder_one', 'birthorder_two', 'birthorder_threemore', 'exposuremedia_no', 'exposuremedia_yes', 'agegroup_younger', 'agegroup_older']

print("n\target columns: {}", format(target_cols))

n arget columns: {} completeCOC

X=drc2[feature_cols]

y=drc2[target_cols]

###describe the outcome

y.value_counts()

incomplete 7850
complete 691

####split into train and test datasets

X_train, X_test, Y_train, y_test=train_test_split(X, y, stratify=y,train_size=.80, random_state=0)

###standard scale of X features
scaler = MinMaxScaler()
scaler.fit(X_train)

X_train = scaler.transform(X_train)
X_test = scaler.transform(X_test)

###assigning unique values to Y
y_test = y_test.map({'incomplete': 1, 'complete': 0}).astype(int)

Y_train = Y_train.map({'incomplete': 1, 'complete': 0}).astype(int)

applying KmeansSMOTE to X and Y train data to correct class imbalance

```
kmeans_smote = KMeansSMOTE(
    kmeans_estimator=KMeans(n_clusters=8, random_state=0),
        cluster_balance_threshold=0.06,
          sampling_strategy='minority',
           k_neighbors=1
```

```
X_train_res, y_train_res = kmeans_smote.fit_resample(X_train, Y_train)
y_train_res.value_counts()
0    6281
1    6279
Name: completeCOC, dtype: int64


#### Train the logistic regression
lg = LogisticRegression(solver='lbfgs',max_iter=300,random_state = 0)
lg.fit(X_train_res, y_train_res)
##Test the trained Logistic regression model
lg_prediction = lg.predict(X_test)
print(accuracy_score(lg_prediction, y_test))
print('Precision: %.3f' % precision_score(lg_prediction, y_test))
print('Recall: %.3f' % recall_score(lg_prediction, y_test))
print('F1 Score: %.3f' % f1_score(lg_prediction, y_test))
confusion_matrix(y_test, lg_prediction)


###Train the Support Vector Machine
svc = SVC(kernel = 'rbf', random_state = 0)
svc.fit(X_train_res, y_train_res)


###Test the model
svc_prediction = svc.predict(X_test)
print(accuracy_score(svc_prediction, y_test))

###hypertune the parameters of the model
param_grid = {'C': [0.1, 1, 10, 100, 1000],
              'gamma': [1, 0.1, 0.01, 0.001, 0.0001],
              'kernel': ['rbf']}
random_state=(0)
grid2 = GridSearchCV(SVC(), param_grid, refit = True, verbose = 3)

# fitting the model for grid search
svm2=grid2.fit(X_train_res, y_train_res)
###print the best parameters
print(svm2.best_params_)

# print how our model looks after hyper-parameter tuning
print(svm2.best_estimator_)
SVC(C=0.1, gamma=1)

###use the best parameters to train the model
svm3 = SVC(kernel = 'rbf', gamma=1, C=1000,probability=True, random_state = 0)
svm3.fit(X_train_res, y_train_res)
```

```
###test the model
svm3_prediction = svm3.predict(X_test)
print(accuracy_score(svm3_prediction, y_test))
print('Precision: %.3f' % precision_score(svm3_prediction, y_test))
print('Recall: %.3f' % recall_score(svm3_prediction, y_test))
print('F1 Score: %.3f' % f1_score(svm3_prediction, y_test))

### neural network model training
nnw = MLPClassifier(solver='sgd', max_iter=300, random_state=0)
nnw.fit(X_train_res, y_train_res)
###test the model
nnw_prediction = nnw.predict(X_test)
print(accuracy_score(nnw_prediction, y_test))

###neural network hyperparameter tuning
mlp_gs = MLPClassifier(max_iter=300)
parameter_space = {
    'activation': ['tanh', 'relu'],
    'solver': ['sgd', 'adam'],
    'alpha': [0.0001, 0.05],
}

random.seed(0)
nnw2= GridSearchCV(mlp_gs, parameter_space, n_jobs=-1, cv=5)
nnw2.fit(X_train_res, y_train_res)

###print the best parameters
print(nnw2.best_params_)
{'activation': 'relu', 'alpha': 0.0001, 'solver': 'sgd'}
# print how our model looks after hyper-parameter tuning
print(nnw2.best_estimator_)
MLPClassifier(alpha=0.05, max_iter=300, solver='sgd')

###train the model with the best parameters
nnw3 = MLPClassifier(solver='adam', alpha=0.0001,
                  max_iter=600, activation='relu', random_state=0)
nnw3.fit(X_train_res, y_train_res)

###test the model
nnw3_prediction = nnw3.predict(X_test)
print(accuracy_score(nnw3_prediction, y_test))
print('Precision: %.3f' % precision_score(nnw3_prediction, y_test))
print('Recall: %.3f' % recall_score(nnw3_prediction, y_test))
print('F1 Score: %.3f' % f1_score(nnw3_prediction, y_test))
```

```
#### decision tree training
dt = DecisionTreeClassifier(random_state = 0)
dt.fit(X_train_res, y_train_res)

***test the model
dt_prediction = dt.predict(X_test)
print(accuracy_score(dt_prediction, y_test))



###hyper parameter tuning the decision tree
params = {
    'criterion':  ['gini', 'entropy'],
    'max_depth':  [None, 2, 4, 6, 8, 10],
    'max_features': [None, 'sqrt', 'log2', 0.2, 0.4, 0.6, 0.8],
    'splitter': ['best', 'random']
}
random.seed(0)
dt2=GridSearchCV(estimator=DecisionTreeClassifier(),param_grid=params,cv=5,verbose=3)

###print best parameters
print(dt2.best_params_)
dt3 = DecisionTreeClassifier(criterion='entropy', max_depth=4, max_features=0.2, splitter='best',
random_state = 0)

###train the model with the best paramters
dt3.fit(X_train_res, y_train_res)

###test the model
dt3_prediction = dt3.predict(X_test)
print(accuracy_score(dt3_prediction, y_test))
print('Precision: %.3f' % precision_score( dt3_prediction,y_test))
print('Recall: %.3f' % recall_score(dt3_prediction, y_test))
print('F1 Score: %.3f' % f1_score(dt3_prediction, y_test))

####train the Random Forest
rf = RandomForestClassifier(random_state = 0)
rf.fit(X_train_res, y_train_res)
###test the model
rf_prediction = rf.predict(X_test)
print(accuracy_score(rf_prediction, y_test))

###hypertuning random forest
# Number of trees in random forest
```

```
n_estimators = [int(x) for x in range(200,2000,200)]
# Number of features to consider at every split
max_features = [None,'log2','sqrt' , 1,10]
# Maximum number of levels in tree
max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
max_depth.append(None)
# Minimum number of samples required to split a node
min_samples_split = [2, 5, 10]
# Minimum number of samples required at each leaf node
min_samples_leaf = [1, 2, 4]
# Method of selecting samples for training each tree
bootstrap = [True, False]
criterion=["gini", "entropy"]
# Create the random grid
# Create the random grid
random_grid = {'n_estimators': n_estimators,
               'max_features': max_features,
               'max_depth': max_depth,
               'min_samples_split': min_samples_split,
               'min_samples_leaf': min_samples_leaf,
               'criterion': criterion,
               'bootstrap': bootstrap}
print(random_grid)
rf_random = RandomizedSearchCV(estimator = rf, param_distributions = random_grid,
 n_iter = 100, cv = 3, verbose=2, random_state=0, n_jobs = -1)


# Fit the random search model
rf_random.fit(X_train_res, y_train_res)


###print the best parameters

rf_random.best_params_


###train the model with the best parameters
rf1 = RandomForestClassifier(n_estimators=1200,min_samples_split=2, min_samples_leaf=4,
                          max_features=1,max_depth=70, bootstrap=True, criterion='gini', random_state = 0)
rf1.fit(X_train_res, y_train_res)


###test the model
rf1_prediction = rf1.predict(X_test)
print(accuracy_score(rf1_prediction, y_test))
print('Precision: %.3f' % precision_score(rf1_prediction, y_test))
print('Recall: %.3f' % recall_score(rf1_prediction, y_test))
print('F1 Score: %.3f' % f1_score(rf1_prediction, y_test))



###PLOR ROC curves
```

```
###logistic regression
lg_pred = lg.predict_proba(X_test)[::,1]
lg_auc = roc_auc_score(y_test, lg_pred)
print('Logistic Regression AUROC = {:.3f}'.format(lg_auc))
lg_fpr, lg_tpr, _ = roc_curve(y_test, lg_pred )

###random forest
rf_pred = rf1.predict_proba(X_test)[::,1]
rf_auc = roc_auc_score(y_test, rf_pred)
print('Random Forest Regression AUROC = {:.3f}'.format(rf_auc))
rf_fpr, rf_tpr, _=metrics.roc_curve(y_test, rf_pred)

####decision tree
dt_pred = dt3.predict_proba(X_test)[::,1]
dt_auc = roc_auc_score(y_test, dt_pred)
print('Decision tree Regression AUROC = {:.3f}'.format(dt_auc))
dt_fpr, dt_tpr, _=roc_curve(y_test, dt_pred)

###support vector machine
svm_pred = svm3.predict_proba(X_test)[::,1]
svm_auc = roc_auc_score(y_test, svm_pred)
print('SVM Regression AUROC = {:.3f}'.format(svm_auc))
svm_fpr, svm_tpr,_=roc_curve(y_test, svm_pred)

###artificial neural network
ann_pred = nnw3.predict_proba(X_test)[::,1]
ann_auc = roc_auc_score(y_test, ann_pred)
print('ANN Regression AUROC = {:.3f}'.format(ann_auc))
ann_fpr, ann_tpr, _=roc_curve(y_test,ann_pred)

plt.plot(lg_fpr, lg_tpr, linestyle = '--', label = 'LR (AUROC ={:.2f})'.format(lg_auc))
plt.plot(rf_fpr, rf_tpr, linestyle = '--', label = 'RF (AUROC ={:.2f})'.format(rf_auc))
plt.plot(dt_fpr, dt_tpr, linestyle = '--', label = 'DT (AUROC ={:.2f})'.format(dt_auc))
plt.plot(svm_fpr, svm_tpr, linestyle = '--', label = 'SVM (AUROC ={:.2f})'.format(svm_auc))
plt.plot(ann_fpr, ann_tpr, linestyle = '--', label = 'ANN(AUROC ={:.2f})'.format(ann_auc))
#Title
#plt.title('ROC Plot')
#Axis labels
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
#Show legend
plt.legend(loc="lower right")

plt.show()
```

```
###feature importance using the Random Forest
perm_importance = permutation_importance(rf1, X_test, y_test, n_repeats=10, random_state=0)
sorted_idx = perm_importance.importances_mean.argsort()
fig = plt.figure(figsize=(12, 6))
plt.barh(range(len(sorted_idx)), perm_importance.importances_mean[sorted_idx], align='center')
plt.yticks(range(len(sorted_idx)), np.array(X.columns)[sorted_idx])
```

REFERENCES

1. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Education Quarterly. 1988;15(4):351-77.
2. UN Inter-agency Group for Child Mortality Estimation. Levels & Trends in Child Mortality. 2018.
3. World Health Organization. WHO recommendations on antenatal care for a positive pregnancy experience: World Health Organization; 2016.
4. World Health Organisation. WHO antenatal care randomized trial: manual for the implementation of the new model. Geneva: World Health Organisation; 2001.
5. World Health Organization. Making pregnancy safer: the critical role of the skilled attendant: a joint statement by WHO, ICM and FIGO: World Health Organization; 2004.
6. World Health Organization. WHO recommendations on postnatal care of the mother and newborn: World Health Organization; 2014.
7. Iqbal S, Maqsood S, Zakar R, Zakar MZ, Fischer F. Continuum of care in maternal, newborn and child health in Pakistan: analysis of trends and determinants from 2006 to 2012. BMC Health Services Research. 2017;17(1):1-15.
8. Alem AZ, Yeshaw Y, Liyew AM, Tesema GA, Alamneh TS, Worku MG, et al. Timely initiation of antenatal care and its associated factors among pregnant women in sub-Saharan Africa: A multicountry analysis of Demographic and Health Surveys. PloS One. 2022;17(1):e0262411.
9. Okedo-Alex IN, Akamike IC, Ezeanosike OB, Uneke CJ. Determinants of antenatal care utilisation in sub-Saharan Africa: a systematic review. BMJ Open. 2019;9(10):e031890.
10. Tessema ZT, Tesema GA. Pooled prevalence and determinants of skilled birth attendant delivery in East Africa countries: a multilevel analysis of Demographic and Health Surveys. Italian journal of pediatrics. 2020;46(1):1-11.
11. Tessema ZT, Yazachew L, Tesema GA, Teshale AB. Determinants of postnatal care utilization in sub-Saharan Africa: a meta and multilevel analysis of data from 36 sub-Saharan countries. Italian journal of pediatrics. 2020;46(1):1-11.
12. Samuel O, Zewotir T, North D. Application of machine learning methods for predicting under-five mortality: analysis of Nigerian demographic health survey 2018 dataset. BMC Medical Informatics and Decision Making. 2024;24(1):86.
13. UNICEF. Neonatal mortality 2021 [Available from: https://data.unicef.org/topic/child-survival/neonatal-mortality.
14. Afolabi BM. Sub-sahara african neonates-ghosts to statistics. J Neonatal Biol. 2017;6(1):246.
15. Lawn JE, Kerber K, Enweronu-Laryea C, Cousens S, editors. 3.6 million neonatal deaths—what is progressing and what is not? Seminars in Perinatology; 2010: Elsevier.
16. Grady SC, Frake AN, Zhang Q, Bene M, Jordan DR, Vertalka J, et al. Neonatal mortality in East Africa and West Africa: a geographic analysis of district-level demographic and health survey data. Geospatial health. 2017;12(1).

17. Tujare Tunta Orsido TTO, Netsanet Abera Asseffa NAA, Tezera Moshago Berheto TMB. Predictors of neonatal mortality in neonatal intensive care unit at Referral Hospital in Southern Ethiopia: a retrospective cohort study. 2019.
18. Tiruye G, Shiferaw K, Shunu A, Sintayeu Y, Seid AM. Antenatal care predicts neonatal mortality in Eastern Africa: a systematic review and meta-analysis of observational studies. Journal of Neonatology. 2022;36(1):42-54.
19. Kawungezi PC, AkiiBua D, Aleni C, Chitayi M, Niwaha A, Kazibwe A, et al. Attendance and utilization of antenatal care (ANC) services: multi-center study in upcountry areas of Uganda. Open journal of preventive medicine. 2015;5(3):132.
20. Gebresilassie B, Belete T, Tilahun W, Berhane B, Gebresilassie S. Timing of first antenatal care attendance and associated factors among pregnant women in public health institutions of Axum town, Tigray, Ethiopia, 2017: a mixed design study. BMC pregnancy and childbirth. 2019;19:1-11.
21. World Health Organisation. WHO recommendations on antenatal care for a positive pregnancy experience: World Health Organization; 2016.
22. Sserwanja Q, Nuwabaine L, Kamara K, Musaba MW. Prevalence and factors associated with utilisation of postnatal care in Sierra Leone: a 2019 national survey. BMC Public Health. 2022;22(1):1-10.
23. World Health Organisation. Maternal mortality: fact sheet: to improve maternal health, barriers that limit access to quality maternal health services must be identified and addressed at all levels of the health system. Geneva: WHO; 2014.
24. Herrera CA, Lewin S, Paulsen E, Ciapponi A, Opiyo N, Pantoja T, et al. Governance arrangements for health systems in low-income countries: an overview of systematic reviews. Cochrane Database of Systematic Reviews. 2017(9).
25. United Nations Department of Econonomic and Social Affairs Sustainable Development. Ensure healthy lives and promote well-being for all at all ages 2023 [cited 2023 11 May]. Available from: https://sdgs.un.org/goals/goal3.
26. Andegiorgish AK, Elhoumed M, Qi Q, Zhu Z, Zeng L. Determinants of antenatal care use in nine sub-Saharan African countries: a statistical analysis of cross-sectional data from Demographic and Health Surveys. BMJ open. 2022;12(2):e051675.
27. Bobo FT, Asante A, Woldie M, Dawson A, Hayen A. Spatial patterns and inequalities in skilled birth attendance and caesarean delivery in sub-Saharan Africa. BMJ Global Health. 2021;6(10):e007074.
28. Tsawe M, Moto A, Netshivhera T, Ralesego L, Nyathi C, Susuman AS. Factors influencing the use of maternal healthcare services and childhood immunization in Swaziland. International journal for equity in health. 2015;14(1):1-11.
29. Nahar MT, Anik SFI, Islam MA, Islam SMS. Individual and community-level factors associated with skilled birth attendants during delivery in Bangladesh: A multilevel analysis of demographic and health surveys. Plos one. 2022;17(6):e0267660.
30. Dickson KS, Ayebeng C, Adu-Gyamfi AB, Okyere J. Postnatal care service utilisation for babies within the first two months after childbirth: an analysis of rural-urban differences in eleven Sub-Saharan African countries. BMC Pregnancy and Childbirth. 2023;23(1):1-10.
31. Sertsewold SG, Debie A, Geberu DM. Continuum of maternal healthcare services utilisation and associated factors among women who gave birth in Siyadebirena Wayu district, Ethiopia: community-based cross-sectional study. BMJ Open. 2021;11(11):e051148.
32. Stewart CL, Hall JA. Factors that affect the utilisation of maternal healthcare in the Mchinji District of Malawi. Plos One. 2022;17(12):e0279613.
33. Gebremedhin AF, Dawson A, Hayen A. Determinants of continuum of care for maternal, newborn, and child health services in Ethiopia: Analysis of the modified composite coverage index using a quantile regression approach. PloS one. 2023;18(1):e0280629.

34. Tesfaye B, Atique S, Azim T, Kebede MM. Predicting skilled delivery service use in Ethiopia: dual application of logistic regression and machine learning algorithms. BMC Medical Informatics and Decision Making. 2019;19(1):1-10.
35. Hammoud B, Semaan A, Elhajj I, Benova L. Can machine learning models predict maternal and newborn healthcare providers' perception of safety during the COVID-19 pandemic? A cross-sectional study of a global online survey. Human Resources for Health. 2022;20(1):63.
36. Imbo AE, Mbuthia EK, Ngotho DN. Determinants of neonatal mortality in Kenya: evidence from the Kenya demographic and health survey 2014. International Journal of Maternal and Child Health and AIDS. 2021;10(2):287.
37. Amouzou A, Ziqi M, Carvajal–Aguirre L, Quinley J. Skilled attendant at birth and newborn survival in Sub–Saharan Africa. Journal of global health. 2017;7(2).
38. Kananura RM, Wamala R, Ekirapa-Kiracho E, Tetui M, Kiwanuka SN, Waiswa P, et al. A structural equation analysis on the relationship between maternal health services utilization and newborn health outcomes: a cross-sectional study in Eastern Uganda. BMC Pregnancy and Childbirth. 2017;17(1):1-12.
39. Kananura RM, Wamala R, Ekirapa-Kiracho E, Tetui M, Kiwanuka SN, Waiswa P, et al. A structural equation analysis on the relationship between maternal health services utilization and newborn health outcomes: a cross-sectional study in Eastern Uganda. BMC pregnancy and childbirth. 2017;17:1-12.
40. Gunzler D, Chen T, Wu P, Zhang H. Introduction to mediation analysis with structural equation modeling. Shanghai archives of psychiatry. 2013;25(6):390.
41. World Health Organisation. Newborn for Africa. Geneva: WHO; 2019.
42. Nguyen TYC, Fagbayigbo BO, Cissé G, Redi N, Fuhrimann S, Okedi J, et al. Diarrhoea among children aged under five years and risk factors in informal settlements: A cross-sectional study in Cape Town, South Africa. International journal of environmental research and public health. 2021;18(11):6043.
43. World Health Orgnaisation. Pneumonia in children 2022 [Available from: https://www.who.int/news-room/fact-sheets/detail/pneumonia.
44. Cheng A, Ghanem-Uzqueda A, Hoff NA, Ashbaugh H, Doshi RH, Mukadi P, et al. Tetanus seroprotection among children in the Democratic Republic of the Congo, 2013–2014. Plos one. 2022;17(5):e0268703.
45. World Health Organisation. Preterm birth 2023 [Available from: https://www.who.int/news-room/fact-sheets/detail/pneumonia.
46. World Health Organisation. Perinatal asphyxia 2012 [Available from: https://www.who.int/teams/maternal-newborn-child-adolescent-health-and-ageing/newborn-health/perinatal-asphyxia.
47. World Health Organisation. Congenital disorders 2021 [Available from: https://www.who.int/health-topics/congenital-anomalies.
48. World Health Organisation. WHO Recommended Interventions for Improving Maternal and Newborn Health 2018 [cited 2023 19 March]. Available from: http://apps.who.int/iris/bitstream/handle/10665/69509/WHO_MPS_07.05_eng.pdf?sequence=1.
49. Nieuwoudt L, Mackay CA, Mda S. Causes of and Modifiable Factors Contributing to Neonatal Deaths at Dora Nginza Hospital in the Eastern Cape, South Africa. Global Pediatric Health. 2022;9:2333794X221139413.
50. Islam MA, Butt ZA, Sathi NJ. Prevalence of Neonatal Mortality and its Associated Factors: A Meta-analysis of Demographic and Health Survey Data from 21 Developing Countries. Dr Sulaiman Al Habib Medical Journal. 2022;4(3):145-52.
51. Kuuire VZ, Kangmennaang J, Atuoye KN, Antabe R, Boamah SA, Vercillo S, et al. Timing and utilisation of antenatal care service in Nigeria and Malawi. Global Public Health. 2017;12(6):711-27.
52. Singh K, Story WT, Moran AC. Assessing the continuum of care pathway for maternal health in South Asia and sub-Saharan Africa. Maternal and Child Health journal. 2016;20:281-9.

53. Yemane GD. The factors associated with antenatal care utilization in Ethiopia. Annals of Medicine and Surgery. 2022;79:104092.
54. Kasagama E, Todd J, Renju J. Factors associated with changes in adequate antenatal care visits among pregnant women aged 15-49 years in Tanzania from 2004 to 2016. BMC Pregnancy and Childbirth. 2022;22(1):18.
55. Marques BL, Tomasi YT, Saraiva SdS, Boing AF, Geremia DS. Guidelines to pregnant women: the importance of the shared care in primary health care. Escola Anna Nery. 2020;25.
56. Kuhnt J, Vollmer S. Antenatal care services and its implications for vital and health outcomes of children: evidence from 193 surveys in 69 low-income and middle-income countries. BMJ Open. 2017;7(11):e017122.
57. Haftu A, Hagos H, Mehari M-A, G/her B. Pregnant women adherence level to antenatal care visit and its effect on perinatal outcome among mothers in Tigray public health institutions, 2017: cohort study. BMC Research Notes. 2018;11:1-6.
58. Negero MG, Sibbritt D, Dawson A. Women's utilisation of quality antenatal care, intrapartum care and postnatal care services in Ethiopia: a population-based study using the demographic and health survey data. BMC Public Health. 2023;23(1):1174.
59. Negero MG, Mitike YB, Worku AG, Abota TL. Skilled delivery service utilization and its association with the establishment of Women's Health Development Army in Yeky district, South West Ethiopia: a multilevel analysis. BMC Research Notes. 2018;11(1):1-9.
60. Seidu A-A, Ahinkorah BO, Agbaglo E, Oduro JK, Amoah A, Yaya S. Factors associated with the utilisation of skilled delivery services in Papua New Guinea: evidence from the 2016–2018 Demographic and Health Survey. International Health. 2022;14(1):64-73.
61. Esopo K, Derby L, Haushofer J. Interventions to improve adherence to antenatal and postnatal care regimens among pregnant women in sub-Saharan Africa: a systematic review. BMC Pregnancy and Childbirth. 2020;20(1):1-12.
62. Sacks E, Finlayson K, Brizuela V, Crossland N, Ziegler D, Sauvé C, et al. Factors that influence uptake of routine postnatal care: Findings on women's perspectives from a qualitative evidence synthesis. PloS One. 2022;17(8):e0270264.
63. Dona A, Tulicha T, Arsicha A, Dabaro D. Factors influencing utilization of early postnatal care services among postpartum women in Yirgalem town, Sidama Regional State, Ethiopia. SAGE Open Medicine. 2022;10:20503121221088098.
64. Haile D, Kondale M, Andarge E, Tunje A, Fikadu T, Boti N. Level of completion along continuum of care for maternal and newborn health services and factors associated with it among women in Arba Minch Zuria woreda, Gamo zone, Southern Ethiopia: A community based cross-sectional study. Plos One. 2020;15(6):e0221670.
65. Kerber KJ, de Graft-Johnson JE, Bhutta ZA, Okong P, Starrs A, Lawn JE. Continuum of care for maternal, newborn, and child health: from slogan to service delivery. The Lancet. 2007;370(9595):1358-69.
66. Buli TD, Wakgari N, Ganfure G, Wondimu F, Dube DL, Moti G, et al. Completion of the continuum of maternity care and associated factors among women who gave birth in the last 6 months in Chelia district, West Shoa zone, Ethiopia: A community-based cross-sectional study. Frontiers in Public Health. 2022;10.
67. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Annals of behavioral medicine. 2003;26:172-81.
68. Zheng D, Hao X, Khan M, Wang L, Li F, Xiang N, et al. Comparison of machine learning and logistic regression as predictive models for adverse maternal and neonatal outcomes of preeclampsia: A retrospective study. Frontiers in Cardiovascular Medicine. 2022;9:959649.
69. Atuhaire R, Atuhaire LK, Wamala R, Nansubuga E. Interrelationships between early antenatal care, health facility delivery and early postnatal care among women in Uganda: a structural equation analysis. Global health action. 2020;13(1):1830463.

70. Oh J, Moon J, Choi JW, Kim K. Factors associated with the continuum of care for maternal, newborn and child health in The Gambia: a cross-sectional study using Demographic and Health Survey 2013. BMJ Open. 2020;10(11):e036516.
71. Bhowmik KR, Das S. On selection of an appropriate logistic model to determine the risk factors of childhood stunting in B angladesh. Maternal & Child Nutrition. 2019;15(1):e12636.
72. Buis ML. Predict and adjust with logistic regression. The Stata Journal. 2007;7(2):221-6.
73. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 2015;349(6245):255-60.
74. Lecca P. Machine learning for causal inference in biological networks: perspectives of this challenge. Frontiers in Bioinformatics. 2021;1:746712.
75. Loring Z, Mehrotra S, Piccini JP. Machine learning in 'big data': handle with care. EP Europace. 2019;21(9):1284-5.
76. Islam MN, Mustafina SN, Mahmud T, Khan NI. Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda. BMC Pregnancy and Childbirth. 2022;22(1):1-19.
77. Sarker IH. Machine learning: Algorithms, real-world applications and research directions. SN computer science. 2021;2(3):160.
78. Fernández L, Mediano P, García R, Rodríguez JM, Marín M. Risk factors predicting infectious lactational mastitis: decision tree approach versus logistic regression analysis. Maternal and child health journal. 2016;20:1895-903.
79. Breiman L. Random forests. Machine Learning. 2001;45:5-32.
80. Gutierrez S, Tardaguila J, Fernandez-Novales J, Diago MP. Support vector machine and artificial neural network models for the classification of grapevine varieties using a portable NIR spectrophotometer. PloS One. 2015;10(11):e0143197.
81. Lei PW, Wu Q. Introduction to structural equation modeling: Issues and practical considerations. Educational Measurement: Issues and Practice. 2007;26(3):33-43.
82. Rabe-Hesketh S, Skrondal A, Pickles A. Generalized multilevel structural equation modeling. Psychometrika. 2004;69:167-90.
83. Schedwin M, Furaha AB, Kapend R, Akilimali P, Malembaka EB, Hildenwall H, et al. Under-five mortality in the Democratic Republic of the Congo: secondary analyses of survey and conflict data by province. Bulletin of the World Health Organization. 2022;100(7):422.
84. Qi X, Zhu Y, Wang Y, He Q, Hee J, Chen W, et al. Socio-economic inequalities in children's nutritional status in Democratic Republic of the Congo in 2017–2018: an analysis of data from a nationally representative survey. Public Health Nutrition. 2022;25(2):257-68.
85. Wang H, Frasco E, Takesue R, Tang K. Maternal education level and maternal healthcare utilization in the Democratic Republic of the Congo: an analysis of the multiple indicator cluster survey 2017/18. BMC Health Services Research. 2021;21(1):1-13.
86. Lang'at E, Mwanri L, Temmerman M. Effects of implementing free maternity service policy in Kenya: an interrupted time series analysis. BMC Health Services Research. 2019;19(1):1-10.
87. Gitobu CM, Gichangi PB, Mwanda WO. The effect of Kenya's free maternal health care policy on the utilization of health facility delivery services and maternal and neonatal mortality in public health facilities. BMC Pregnancy and Childbirth. 2018;18(1):1-11.
88. Kiplagat IJ. Determinants of health insurance choice in Kenya: University of Nairobi, Kenya; 2011.
89. Mangu CD, Rumisha SF, Lyimo EP, Mremi IR, Massawe IS, Bwana VM, et al. Trends, patterns and cause-specific neonatal mortality in Tanzania: a hospital-based retrospective survey. International Health. 2021;13(4):334-43.
90. Kruk ME, Mbaruku G, Rockers PC, Galea S. User fee exemptions are not enough: out-of-pocket payments for 'free'delivery services in rural Tanzania. Tropical Medicine & International Health. 2008;13(12):1442-51.

91. Kapologwe NA, Kagaruki GB, Kalolo A, Ally M, Shao A, Meshack M, et al. Barriers and facilitators to enrollment and re-enrollment into the community health funds/Tiba Kwa Kadi (CHF/TIKA) in Tanzania: a cross-sectional inquiry on the effects of socio-demographic factors and social marketing strategies. BMC Health Services Research. 2017;17:1-9.
92. Alamneh A, Asmamaw A, Woldemariam M, Yenew C, Atikilt G, Andualem M, et al. Trend change in delayed first antenatal care visit among reproductive-aged women in Ethiopia: multivariate decomposition analysis. Reproductive Health. 2022;19(1):1-13.
93. Gupta S, Yamada G, Mpembeni R, Frumence G, Callaghan-Koru JA, Stevenson R, et al. Factors associated with four or more antenatal care visits and its decline among pregnant women in Tanzania between 1999 and 2010. PloS One. 2014;9(7):e101893.
94. Sampurna MTA, Handayani KD, Utomo MT, Angelika D, Etika R, Harianto A, et al. Determinants of neonatal deaths in Indonesia: A national survey data analysis of 10,838 newborns. Heliyon. 2023;9(1).
95. Ruktanonchai CW, Nilsen K, Alegana VA, Bosco C, Ayiko R, Seven Kajeguka AC, et al. Temporal trends in spatial inequalities of maternal and newborn health services among four east African countries, 1999–2015. BMC Public Health. 2018;18:1-13.
96. East African Community. The Democratic Republic of the Congo joins EAC as its 7th Member 2022 [Available from: https://www.eac.int/press-releases/2402-the-democratic-republic-of-the-congo-joins-eac-as-its-7th-member.
97. The World Bank & World Development. The World Bank in DRC 2021 [Available from: https://www.worldbank.org/en/country/drc/overview.
98. One Planet Nations Online. Administrative Map of Democratic Republic of the Congo 2023 [Available from: https://www.nationsonline.org/oneworld/map/dr_congo_map.htm.
99. World Bank. The World Bank in the DRC 2023 [Available from: https://www.worldbank.org/en/country/drc/overview.
100. Kenya National Bureau of Statistics MoHK NACCK, Kenya Medical Research Institute, Population NCf, Development/Kenya,. The 2019 Kenya Population and Housing Census: Population by County and Sub-county. Kenya; 2019.
101. World Bank. The World Bank in Kenya 2023 [Available from: https://www.worldbank.org/en/country/kenya/overview.
102. Macharia PM, Giorgi E, Thuranira PN, Joseph NK, Sartorius B, Snow RW, et al. Sub national variation and inequalities in under-five mortality in Kenya since 1965. BMC Public Health. 2019;19(1):1-12.
103. UNESCO Institute for Statistics. United Republic of Tanzania 2019 2019 [Available from: http://uis.unesco.org/country/TZ.
104. One Planet Nations Online. Administrative Map of Tanzania 2023 [Available from: https://www.nationsonline.org/oneworld/map/tanzania-administrative-map.
105. Bank W. Poverty & Equity Brief Africa Eastern & Southern :Tanzania. World Bank; 2021.
106. USAID. The DHS Program 2023 [cited 2023 1 June]. Available from: https://dhsprogram.com/data.
107. The DHS Program. DHS Model Questionnaire - Phase 7 (English, French) 2024 [Available from: https://dhsprogram.com/publications/publication-dhsq7-dhs-questionnaires-and-manuals.cfm.
108. Nigatu SG, Birhan TY. The magnitude and determinants of delayed initiation of antenatal care among pregnant women in Gambia; evidence from Gambia demographic and health survey data. BMC Public Health. 2023;23(1):1-10.
109. Mgata S, Maluka SO. Factors for late initiation of antenatal care in Dar es Salaam, Tanzania: A qualitative study. BMC Pregnancy and Childbirth. 2019;19:1-9.
110. Paudel YR, Jha T, Mehata S. Timing of first antenatal care (ANC) and inequalities in early initiation of ANC in Nepal. Frontiers in public health. 2017;5:242.
111. StataCorp L. Stata Statistical Software: Release 17 College Station, TX: StataCorp LLC2021 [

112. ArcGIS. ArcGIS Release 10.1 Redlands, USA2012 [
113. Amsalu G, Talie A, Gezimu W, Duguma A. Non-utilization of postnatal care and its associated factors among women who gave birth in rural districts of Northern Ethiopia: A community-based mixed-method study. Women's Health. 2022;18:17455057221125091.
114. Kotsiantis S. Feature selection for machine learning classification problems: a recent overview. Artificial intelligence review. 2011;42(1):157-76.
115. Kulkarni A, Chong D, Batarseh FA. Foundations of data imbalance and solutions for a data democracy. Data democracy: Elsevier; 2020. p. 83-106.
116. Song Y-Y, Ying L. Decision tree methods: applications for classification and prediction. Shanghai archives of psychiatry. 2015;27(2):130.
117. Shreffler J, Huecker MR. Diagnostic testing accuracy: Sensitivity, specificity, predictive values and likelihood ratios. 2020.
118. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. Journal of Thoracic Oncology. 2010;5(9):1315-6.
119. Yang S, Berdine G. The receiver operating characteristic (ROC) curve. The Southwest Respiratory and Critical Care Chronicles. 2017;5(19):34-6.
120. Tadese M, Tessema SD, Aklilu D, Wake GE, Mulu GB. Dropout from a maternal and newborn continuum of care after antenatal care booking and its associated factors in Debre Berhan Town, Northeast Ethiopia. Frontiers in Medicine.2865.
121. Zhang J, Li Z, Pu Z, Xu C. Comparing prediction performance for crash injury severity among various machine learning and statistical methods. IEEE Access. 2018;6:60079-87.
122. Ozcan I, Aydin H, Cetinkaya A. Comparison of Classification Success Rates of Different Machine Learning Algorithms in the Diagnosis of Breast Cancer. Asian Pacific Journal of Cancer Prevention. 2022;23(10):3287-97.
123. Thejas G, Hariprasad Y, Iyengar S, Sunitha N, Badrinath P, Chennupati S. An extension of Synthetic Minority Oversampling Technique based on Kalman filter for imbalanced datasets. Machine Learning with Applications. 2022;8:100267.
124. Towards Data Sceince. Accuracy, Precision, Recall or F1 2023 [Available from: https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9.
125. Bestami Yuksel B, Bahtiyar S, Yilmazer A, editors. Credit Card Fraud Detection with NCA Dimensionality Reduction. 13th International Conference on Security of Information and Networks; 2020.
126. Trevethan R. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. Frontiers in Public Health. 2017;5:307.
127. Fisher G. Endogenous and exogenous investment in macro-economic models. The Review of Economics and Statistics. 1953:211-20.
128. Ho W. Recommended definition, terminology and format for statistical tables related to the perinatal period. Acta Obstet Gynecol Scand. 1997;56:247-53.
129. World Health Organisation. Low birth weight 2024 [Available from: https://www.who.int/data/nutrition/nlis/info/low-birth-weight.
130. Huber C, editor Generalized structural equation modeling using stata. Italian Stata Users Group Meeting, November; 2013.
131. Barmar S, Alimohammadian M, Sadjadi SA, Poustchi H, Hosseini SM, Yasseri M. Generalized Structural Equation Modeling (GSEM) and its application in health researches. 2018.
132. Chakrabarti A, Ghosh JK. AIC, BIC and recent advances in model selection. Philosophy of statistics. 2011:583-605.
133. Jung SJ. Introduction to mediation analysis and examples of its application to real-world data. Journal of Preventive Medicine and Public Health. 2021;54(3):166.
134. Aguinis H, Edwards JR, Bradley KJ. Improving our understanding of moderation and mediation in strategic management research. Organizational Research Methods. 2017;20(4):665-85.

135. VanderWeele T, Vansteelandt S. Mediation analysis with multiple mediators. Epidemiologic Methods. 2014;2(1):95-115.
136. UNDP. SUSTAINABLE DEVELOPMENT GOALS 2021 [Available from: https://www.africa.undp.org/content/rba/en/home/sustainable-development-goals.html.
137. World Health Organization. Trends in maternal mortality: 1990-2015: estimates from WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division: World Health Organization; 2015.
138. Wolde HF, Tsegaye AT, Sisay MM. Late initiation of antenatal care and associated factors among pregnant women in Addis Zemen primary hospital, South Gondar, Ethiopia. Reproductive Health. 2019;16(1):1-8.
139. Ebonwu J, Mumbauer A, Uys M, Wainberg ML, Medina-Marino A. Determinants of late antenatal care presentation in rural and peri-urban communities in South Africa: A cross-sectional study. PLoS One. 2018;13(3):e0191903.
140. East African Community. The Democratic Republic of the Congo joins EAC as its 7th Member 2022 [Available from: https://www.eac.int/press-releases/2402-the-democratic-republic-of-the-congo-joins-eac-as-its-7th-member.
141. Ruktanonchai CW, Nilsen K, Alegana VA, Bosco C, Ayiko R, Seven Kajeguka AC, et al. Temporal trends in spatial inequalities of maternal and newborn health services among four east African countries, 1999–2015. BMC Public Health. 2018;18(1):1-13.
142. USAID. The DHS Program 2022 [cited 2022 7 January]. Available from: https://dhsprogram.com/data/.
143. Rwabilimbo AG, Ahmed KY, Page A, Ogbo FA. Trends and factors associated with the utilisation of antenatal care services during the millennium development goals era in Tanzania. Tropical Medicine and Health. 2020;48(1):1-16.
144. StataCorp L. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC; 2019.
145. ArcGIS Release 10.1 [Internet]. Redlands, CA. 2012.
146. Tesfaye G, Loxton D, Chojenta C, Semahegn A, Smith R. Delayed initiation of antenatal care and associated factors in Ethiopia: a systematic review and meta-analysis. Reproductive Health. 2017;14(1):1-17.
147. Teshale AB, Tesema GA. Prevalence and associated factors of delayed first antenatal care booking among reproductive age women in Ethiopia; a multilevel analysis of EDHS 2016 data. PloS One. 2020;15(7):e0235538.
148. Clark J, Sweet L, Nyoni S, Ward PR. Improving male involvement in antenatal care in low and middle-income countries to prevent mother to child transmission of HIV: a realist review. PloS One. 2020;15(10):e0240087.
149. Dominicus DA, Akamatsu T. Health policy and implementations in Tanzania. The Keio journal of medicine. 1989;38(2):192-200.
150. Dewau R, Muche A, Fentaw Z, Yalew M, Bitew G, Amsalu ET, et al. Time to initiation of antenatal care and its predictors among pregnant women in Ethiopia: Cox-gamma shared frailty model. PloS One. 2021;16(2):e0246349.
151. Mbuagbaw L, Medley N, Darzi AJ, Richardson M, Garga KH, Ongolo-Zogo P. Health system and community level interventions for improving antenatal care coverage and health outcomes. Cochrane Database of Systematic Reviews. 2015(12).
152. Morestin F, Ridde V. The abolition of user fees for health services in Africa Lessons from the literature. Montreal: University of Montreal. 2009.
153. McKinnon B, Harper S, Kaufman JS, Bergevin Y. Removing user fees for facility-based delivery services: a difference-in-differences evaluation from ten sub-Saharan African countries. Health policy and planning. 2015;30(4):432-41.
154. Sinyange N, Sitali L, Jacobs C, Musonda P, Michelo C. Factors associated with late antenatal care booking: population based observations from the 2007 Zambia demographic and health survey. The Pan African Medical Journal. 2016;25.

155. Ahinkorah BO, Seidu A-A, Budu E, Mohammed A, Adu C, Agbaglo E, et al. Factors associated with the number and timing of antenatal care visits among married women in Cameroon: evidence from the 2018 Cameroon Demographic and Health Survey. Journal of Biosocial Science. 2021:1-11.
156. Bbaale E. Factors influencing timing and frequency of antenatal care in Uganda. The Australasian Medical Journal. 2011;4(8):431.
157. Finlayson K, Downe S. Why do women not use antenatal services in low-and middle-income countries? A meta-synthesis of qualitative studies. PLoS Medicine. 2013;10(1):e1001373.
158. Pell C, Menaca A, Were F, Afrah NA, Chatio S, Manda-Taylor L, et al. Factors affecting antenatal care attendance: results from qualitative studies in Ghana, Kenya and Malawi. PloS One. 2013;8(1):e53747.
159. Phillips E, Stoltzfus RJ, Michaud L, Pierre GLF, Vermeylen F, Pelletier D. Do mobile clinics provide high-quality antenatal care? A comparison of care delivery, knowledge outcomes and perception of quality of care between fixed and mobile clinics in central Haiti. BMC Pregnancy and Childbirth. 2017;17(1):1-11.
160. Izudi J, Akwang GD, Amongin D. Early postnatal care use by postpartum mothers in Mundri East County, South Sudan. BMC health services research. 2017;17(1):1-8.
161. Dol J, Hughes B, Bonet M, Dorey R, Dorling J, Grant A, et al. Timing of neonatal mortality and severe morbidity during the postnatal period: a systematic review. JBI Evidence Synthesis. 2023;21(1):98-199.
162. Khaki J. Factors associated with the utilization of postnatal care services among Malawian women. Malawi Medical Journal. 2019;31(1):2-11.
163. Ahmed M, Demissie M, Worku A, Abrha A, Berhane Y. Socio-cultural factors favoring home delivery in Afar pastoral community, northeast Ethiopia: a qualitative study. Reproductive health. 2019;16(1):1-9.
164. Mienye ID, Jere N. A Survey of Decision Trees: Concepts, Algorithms, and Applications. IEEE Access. 2024.
165. Kanté AM, Chung CE, Larsen AM, Exavery A, Tani K, Phillips JF. Factors associated with compliance with the recommended frequency of postnatal care services in three rural districts of Tanzania. BMC pregnancy and childbirth. 2015;15:1-10.
166. Mfateneza E, Rutayisire PC, Biracyaza E, Musafiri S, Mpabuka WG. Application of machine learning methods for predicting infant mortality in Rwanda: analysis of Rwanda demographic health survey 2014–15 dataset. BMC Pregnancy and Childbirth. 2022;22(1):388.
167. Emiru AA, Alene GD, Debelew GT. Women's retention on the continuum of maternal care pathway in west Gojjam zone, Ethiopia: multilevel analysis. BMC pregnancy and childbirth. 2020;20:1-14.
168. Maketa V, Vuna M, Baloji S, Lubanza S, Hendrickx D, Inocêncio da Luz RA, et al. Perceptions of health, health care and community-oriented health interventions in poor urban communities of Kinshasa, Democratic Republic of Congo. PLoS One. 2013;8(12):e84314.
169. Garces A, McClure EM, Chomba E, Patel A, Pasha O, Tshefu A, et al. Home birth attendants in low income countries: who are they and what do they do? BMC pregnancy and childbirth. 2012;12:1-9.
170. World Bank Blogs. Obstacles and opportunities for women's economic empowerment in the DRC 2023 [Available from: https://blogs.worldbank.org/africacan/obstacles-and-opportunities-womens-economic-empowerment-drc.
171. Grand-Guillaume-Perrenoud JA, Origlia P, Cignacco E. Barriers and facilitators of maternal healthcare utilisation in the perinatal period among women with social disadvantage: a theory-guided systematic review. Midwifery. 2022;105:103237.
172. Stasse S, Vita D, Kimfuta J, Da Silveira VC, Bossyns P, Criel B. Improving financial access to health care in the Kisantu district in the Democratic Republic of Congo: acting upon complexity. Global health action. 2015;8(1):25480.

173. Kunkel M, Marete I, Cheng ER, Bucher S, Liechty E, Esamai F, et al., editors. Place of delivery and perinatal mortality in Kenya. Seminars in perinatology; 2019: Elsevier.
174. Moindi RO, Ngari MM, Nyambati VC, Mbakaya C. Why mothers still deliver at home: understanding factors associated with home deliveries and cultural practices in rural coastal Kenya, a cross-section study. BMC public health. 2015;16:1-8.
175. Business Daily. Unemployment among women two times higher than men 2023 [Available from: https://www.businessdailyafrica.com/bd/economy/unemployment-among-women-two-times-higher-than-men--4294416.
176. Tanou M, Kishida T, Kamiya Y. The effects of geographical accessibility to health facilities on antenatal care and delivery services utilization in Benin: a cross-sectional study. Reproductive health. 2021;18:1-11.
177. Ahinkorah BO, Seidu A-A, Budu E, Armah-Ansah EK, Agbaglo E, Adu C, et al. Factors Associated with utilization of postnatal care services in Mali, West Africa. International Journal of Translational Medical Research and Public Health. 2021;5(1):23-32.
178. Aboagye RG, Seidu A-A, Ahinkorah BO, Cadri A, Frimpong JB, Hagan JE, et al. Association between frequency of mass media exposure and maternal health care service utilization among women in sub-Saharan Africa: Implications for tailored health communication and education. Plos one. 2022;17(9):e0275202.
179. ICF International. Tanzania Demographic and Health Survey and Malaria Indicator Survey (DHS-MIS) 2015–2016. 2016.
180. Muhunzi S, Ngocho JS, Mwanamsangu A, Sanga L, Hiza H, Msuya SE, et al. Prevalence, predictors and reasons for home delivery amongst women of childbearing age in Dodoma Municipality in central Tanzania. African Health Sciences. 2020;20(4):1933-42.
181. Ministry of Health and Social Welfare (MOHSW) [Tanzania]. Health Sector Strategic Plan July 2015 –June 2020: Reaching all Households with Quality Health Care. 2016.
182. Massenga J, Jeremiah K, Kitinya W, Kim Y-M, van Roosmalen J, van den Akker T. Receiving antenatal care components and associated factors in Northwestern Tanzania. Plos one. 2023;18(4):e0284049.
183. Konje ET, Msuya IE, Matovelo D, Basinda N, Dewey D. Provision of inadequate information on postnatal care and services during antenatal visits in Busega, Northwest Tanzania: a simulated client study. BMC Health Services Research. 2022;22(1):700.
184. Raatikainen K, Heiskanen N, Heinonen S. Under-attending free antenatal care is associated with adverse pregnancy outcomes. BMC public health. 2007;7:1-8.
185. Khan MN, Harris ML, Loxton D. Assessing the effect of pregnancy intention at conception on the continuum of care in maternal healthcare services use in Bangladesh: Evidence from a nationally representative cross-sectional survey. PloS one. 2020;15(11):e0242729.
186. Khan MN, Harris ML, Oldmeadow C, Loxton D. Effect of unintended pregnancy on skilled antenatal care uptake in Bangladesh: analysis of national survey data. Archives of Public Health. 2020;78(1):1-13.
187. Ayalew HG, Liyew AM, Tessema ZT, Worku MG, Tesema GA, Alamneh TS, et al. Prevalence and factors associated with unintended pregnancy among adolescent girls and young women in sub-Saharan Africa, a multilevel analysis. BMC Women's Health. 2022;22(1):1-9.
188. Kothavale A, Meher T. Level of completion along continuum of care for maternal, newborn and child health services and factors associated with it among women in India: a population-based cross-sectional study. BMC pregnancy and childbirth. 2021;21(1):1-12.
189. UNDP. The SDGS IN ACTION 2022 [cited 2022 02 February]. Available from: https://www.undp.org/library/sdgs-action?c_src=CENTRAL&c_src2=GSR.
190. Wealth Health Organisation. Trends in maternal mortality 2000 to 2017: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. 2019. Report No.: 9241516488.

191. UN Inter-agency Group for Child Mortality Estimation. Levels & Trends in Child Mortality. 2018.
192. World Health Organisation. Newborns: improving survival and well-being 2020 [cited 2022 26 May]. Available from: https://www.who.int/news-room/fact-sheets/detail/newborns-reducing-mortality.
193. World Health Organisation. Maternal mortality 2019 [cited 2019 19 September]. Available from: http://www.who.int/news-room/fact-sheets/detail/maternal-mortality.
194. Yeji F, Shibanuma A, Oduro A, Debpuur C, Kikuchi K, Owusu-Agei S, et al. Continuum of care in a maternal, newborn and child health program in Ghana: low completion rate and multiple obstacle factors. PloS one. 2015;10(12):e0142849.
195. Kanté AM, Chung CE, Larsen AM, Exavery A, Tani K, Phillips JF. Factors associated with compliance with the recommended frequency of postnatal care services in three rural districts of Tanzania. BMC pregnancy and childbirth. 2015;15(1):1-10.
196. World Health Organization. WHO recommendations on maternal and newborn care for a positive postnatal experience. Geneva2022.
197. Mohan D, LeFevre AE, George A, Mpembeni R, Bazant E, Rusibamayila N, et al. Analysis of dropout across the continuum of maternal health care in Tanzania: findings from a cross-sectional household survey. Health policy and planning. 2017;32(6):791-9.
198. Mwangi W, Gachuno O, Desai M, Obor D, Were V, Odhiambo F, et al. Uptake of skilled attendance along the continuum of care in rural Western Kenya: selected analysis from Global Health initiative survey-2012. BMC pregnancy and childbirth. 2018;18(1):1-13.
199. Owili PO, Muga MA, Chou Y-J, Hsu Y-HE, Huang N, Chien L-Y. Associations in the continuum of care for maternal, newborn and child health: a population-based study of 12 sub-Saharan Africa countries. BMC public health. 2016;16(1):1-15.
200. Mfateneza E, Rutayisire PC, Biracyaza E, Musafiri S, Mpabuka WG. Application of machine learning methods for predicting infant mortality in Rwanda: analysis of Rwanda demographic health survey 2014–15 dataset. BMC Pregnancy and Childbirth. 2022;22(1):1-13.
201. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. BMC bioinformatics. 2013;14(1):1-16.
202. Shibanuma A, Yeji F, Okawa S, Mahama E, Kikuchi K, Narh C, et al. The coverage of continuum of care in maternal, newborn and child health: a cross-sectional study of woman-child pairs in Ghana. BMJ global health. 2018;3(4):e000786.
203. Guo F, Qi X, Xiong H, He Q, Zhang T, Zou S, et al. Trends of maternal health service coverage in the Democratic Republic of the Congo: a pooled cross-sectional study of MICS 2010 to 2018. BMC pregnancy and childbirth. 2021;21(1):1-10.
204. Tola W, Negash E, Sileshi T, Wakgari N. Late initiation of antenatal care and associated factors among pregnant women attending antenatal clinic of Ilu Ababor Zone, southwest Ethiopia: A cross-sectional study. PloS one. 2021;16(1):e0246230.
205. Mahiti GR, Mkoka DA, Kiwara AD, Mbekenga CK, Hurtig A-K, Goicolea I. Women's perceptions of antenatal, delivery, and postpartum services in rural Tanzania. Global health action. 2015;8(1):28567.
206. Meda IB, Baguiya A, Ridde V, Ouédraogo HG, Dumont A, Kouanda S. Out-of-pocket payments in the context of a free maternal health care policy in Burkina Faso: a national cross-sectional survey. Health Economics Review. 2019;9(1):1-14.
207. Akowuah JA, Agyei-Baffour P, Awunyo-Vitor D. Determinants of antenatal healthcare utilisation by pregnant women in third trimester in peri-urban Ghana. Journal of tropical medicine. 2018;2018.
208. Akunga D, Menya D, Kabue M. Determinants of postnatal care use in Kenya. African Population Studies. 2014;28(3):1447-59.
209. Abbas AA, Walker GJ. Determinants of the utilization of maternal and child health services in Jordan. International Journal of Epidemiology. 1986;15(3):404-7.

210. Gabrysch S, Campbell OM. Still too far to walk: literature review of the determinants of delivery service use. BMC pregnancy and childbirth. 2009;9(1):1-18.
211. Wairoto KG, Joseph NK, Macharia PM, Okiro EA. Determinants of subnational disparities in antenatal care utilisation: a spatial analysis of demographic and health survey data in Kenya. BMC health services research. 2020;20(1):1-12.
212. Orwa J, Mantel M, Mugerwa M, Brownie S, Pallangyo ES, Mwasha L, et al. Maternal healthcare services use in Mwanza Region, Tanzania: a cross-sectional baseline survey. BMC pregnancy and childbirth. 2019;19(1):1-11.
213. Kapologwe NA, Meara JG, Kengia JT, Sonda Y, Gwajima D, Alidina S, et al. Development and upgrading of public primary healthcare facilities with essential surgical services infrastructure: a strategy towards achieving universal health coverage in Tanzania. BMC Health Services Research. 2020;20(1):1-14.
214. Akinyemi JO, Afolabi RF, Awolude OA. Patterns and determinants of dropout from maternity care continuum in Nigeria. BMC pregnancy and childbirth. 2016;16(1):1-11.
215. Toroitich AM, Dunford L, Armitage R, Tanna S. Patients Access to Medicines–A Critical Review of the Healthcare System in Kenya. Risk Management and Healthcare Policy. 2022;15:361.
216. Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: A scoping review. PloS one. 2019;14(2):e0212356.
217. Pan I, Nolan LB, Brown RR, Khan R, van der Boor P, Harris DG, et al. Machine learning for social services: a study of prenatal case management in Illinois. American journal of public health. 2017;107(6):938-44.
218. Lawn JE, Manandhar A, Haws RA, Darmstadt GL. Reducing one million child deaths from birth asphyxia–a survey of health systems gaps and priorities. Health Research Policy and Systems. 2007;5(1):1-10.
219. Mabaso MH, Ndaba T, Mkhize-Kwitshana ZL. Overview of maternal, neonatal and child deaths in South Africa: challenges, opportunities, progress and future prospects. International Journal of MCH and AIDS. 2014;2(2):182.
220. Tolossa T, Fekadu G, Mengist B, Mulisa D, Fetensa G, Bekele D. Impact of antenatal care on neonatal mortality among neonates in Ethiopia: a systematic review and meta-analysis. Archives of Public Health. 2020;78(1):1-11.
221. Yaya Y, Eide KT, Norheim OF, Lindtjørn B. Maternal and neonatal mortality in south-west Ethiopia: estimates and socio-economic inequality. PloS one. 2014;9(4):e96294.
222. Islam MA, Biswas B. Socio-economic factors associated with increased neonatal mortality: A mixed-method study of Bangladesh and 20 other developing countries based on demographic and health survey data. Clinical Epidemiology and Global Health. 2021;11:100801.
223. Institute of Medicine. Improving birth outcomes: meeting the challenge in the developing world. 2003.
224. Raru TB, Ayana GM, Zakaria HF, Merga BT. Association of higher educational attainment on antenatal care utilization among pregnant women in east africa using Demographic and Health Surveys (DHS) from 2010 to 2018: a multilevel analysis. International journal of women's health. 2022:67-77.
225. Kassar SB, Melo AM, Coutinho SB, Lima MC, Lira PI. Determinants of neonatal death with emphasis on health care during pregnancy, childbirth and reproductive history. Jornal de Pediatria. 2013;89(3):269-77.
226. Mwebesa E, Kagaayi J, Ssebagereka A, Nakafeero M, Ssenkusu JM, Guwatudde D, et al. Effect of four or more antenatal care visits on facility delivery and early postnatal care services utilization in Uganda: a propensity score matched analysis. BMC Pregnancy and Childbirth. 2022;22:1-9.
227. Dahiru T, Oche OM. Determinants of antenatal care, institutional delivery and postnatal care services utilization in Nigeria. Pan African medical journal. 2015;22(1).

228. Fall CH. Fetal malnutrition and long-term outcomes. Maternal and Child Nutrition: The First 1,000 Days. 74: Karger Publishers; 2013. p. 11-25.
229. McCormick MC. The contribution of low birth weight to infant mortality and childhood morbidity. The New England journal of medicine. 1985;312(2):82-90.
230. Arsyi M, Besral B, Herdayati M, Phalkey R. Antenatal Care Services and Incidence of Low Birth Weight: A Comparison of Demographic and Health Surveys in 4 ASEAN Countries. Journal of Preventive Medicine and Public Health. 2022;55(6):559.
231. Duodu PA, Bayuo J, Mensah JA, Aduse-Poku L, Arthur-Holmes F, Dzomeku VM, et al. Trends in antenatal care visits and associated factors in Ghana from 2006 to 2018. BMC pregnancy and childbirth. 2022;22(1):59.
232. Simkhada B, Teijlingen ERv, Porter M, Simkhada P. Factors affecting the utilization of antenatal care in developing countries: systematic review of the literature. Journal of advanced nursing. 2008;61(3):244-60.
233. Toroitich AM, Dunford L, Armitage R, Tanna S. Patients Access to Medicines–A Critical Review of the Healthcare System in Kenya. Risk Management and Healthcare Policy. 2022:361-74.
234. Idris IB, Hamis AA, Bukhori ABM, Hoong DCC, Yusop H, Shaharuddin MA-A, et al. Women’s autonomy in healthcare decision making: a systematic review. BMC women's health. 2023;23(1):643.
235. Omer S, Zakar R, Zakar MZ, Fischer F. The influence of social and cultural practices on maternal mortality: a qualitative study from South Punjab, Pakistan. Reproductive health. 2021;18(1):97.
236. Chimatiro CS, Hajison P, Chipeta E, Muula AS. Understanding barriers preventing pregnant women from starting antenatal clinic in the first trimester of pregnancy in Ntcheu District-Malawi. Reproductive health. 2018;15:1-7.
237. Comfort AB, El Ayadi AM, Camlin CS, Tsai AC, Nalubwama H, Byamugisha J, et al. The role of informational support from women’s social networks on antenatal care initiation: qualitative evidence from pregnant women in Uganda. BMC pregnancy and childbirth. 2022;22(1):708.
238. Divala P. Community perceptions of maternal health services in rural areas in Dedza district in Malawi 2020.
239. Byrne A, Caulfield T, Onyo P, Nyagero J, Morgan A, Nduba J, et al. Community and provider perceptions of traditional and skilled birth attendants providing maternal health care for pastoralist communities in Kenya: a qualitative study. BMC pregnancy and childbirth. 2016;16:1-12.
240. Dahab R, Sakellariou D. Barriers to accessing maternal care in low income countries in Africa: a systematic review. International journal of environmental research and public health. 2020;17(12):4292.
241. Tumuhimbise W, Atukunda EC, Ayebaza S, Katusiime J, Mugyenyi G, Pinkwart N, et al. Maternal health-related barriers and the potentials of mobile health technologies: qualitative findings from a pilot randomized controlled trial in rural southwestern Uganda. Journal of family medicine and primary care. 2020;9(7):3657-62.
242. Sultana A. Factors effect on women autonomy and decision-making power within the household in rural communities. Journal of Applied Sciences Research. 2011;7(1):18-22.
243. RHIhub. Rural Health Information 2024 [Available from: https://www.ruralhealthinfo.org/topics/healthcare-access.
244. Gabrysch S, Cousens S, Cox J, Campbell OM. The influence of distance and level of care on delivery place in rural Zambia: a study of linked national data in a geographic information system. PLoS medicine. 2011;8(1):e1000394.
245. Tanou M, Kamiya Y. Assessing the impact of geographical access to health facilities on maternal healthcare utilization: evidence from the Burkina Faso demographic and health survey 2010. BMC public health. 2019;19:1-8.

246. Mwangakala HA. Accessibility of maternal health information and its influence on maternal health preferences in rural Tanzania: A case study of Chamwino District. South African Journal of Information Management. 2021;23(1):1-9.
247. Mrisho M, Schellenberg JA, Mushi AK, Obrist B, Mshinda H, Tanner M, et al. Understanding home-based neonatal care practice in rural southern Tanzania. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2008;102(7):669-78.
248. Ayele BG, Woldu MA, Gebrehiwot HW, Gebre-Egziabher EG, Gebretnsae H, Hadgu T, et al. Magnitude and determinants for place of postnatal care utilization among mothers who delivered at home in Ethiopia: a multinomial analysis from the 2016 Ethiopian demographic health survey. Reproductive health. 2019;16:1-10.
249. Moshi FV, Mbotwa CH. Determinants for choice of home birth over health facility birth among women of reproductive age in Tanzania: an analysis of data from the 2015-16 Tanzania demographic and health survey and malaria indicator survey. BMC Pregnancy and Childbirth. 2020;20:1-9.
250. Fantaye AW, Gunawardena N, Yaya S. Preferences for formal and traditional sources of childbirth and postnatal care among women in rural Africa: A systematic review. PLoS One. 2019;14(9):e0222110.
251. Mrisho M, Obrist B, Schellenberg JA, Haws RA, Mushi AK, Mshinda H, et al. The use of antenatal and postnatal care: perspectives and experiences of women and health care providers in rural southern Tanzania. BMC pregnancy and childbirth. 2009;9:1-12.
252. Gostin LO, Levit LA, Nass SJ. Beyond the HIPAA privacy rule: enhancing privacy, improving health through research. 2009.
253. Eskandari M, Abbaszadeh A, Borhani F. Barriers of referral system to health care provision in rural societies in Iran. Journal of caring sciences. 2013;2(3):229.
254. Jalali FS, Bikineh P, Delavari S. Strategies for reducing out of pocket payments in the health system: a scoping review. Cost Effectiveness and Resource Allocation. 2021;19(1):1-22.
255. Impwii DK, Kivuti-Bitok L. System barriers to the provision of quality maternal health care in two regional teaching and referral hospitals in Kenya: a qualitative study. Global Health Journal. 2023;7(4):200-5.
256. Nuamah GB, Agyei-Baffour P, Mensah KA, Boateng D, Quansah DY, Dobin D, et al. Access and utilization of maternal healthcare in a rural district in the forest belt of Ghana. BMC pregnancy and childbirth. 2019;19:1-11.
257. Weitzman A. The effects of women's education on maternal health: Evidence from Peru. Social science & medicine. 2017;180:1-9.
258. Sokey PP, Adjei E, Ankrah E. Media use for health information dissemination to rural communities by the Ghana Health Service. Journal of Information Science, Systems and Technology. 2018;2(1):1-18.
259. Fatema K, Lariscy JT. Mass media exposure and maternal healthcare utilization in South Asia. SSM-Population Health. 2020;11:100614.
260. Bain LE, Aboagye RG, Dowou RK, Kongnyuy EJ, Memiah P, Amu H. Prevalence and determinants of maternal healthcare utilisation among young women in sub-Saharan Africa: cross-sectional analyses of demographic and health survey data. BMC public health. 2022;22(1):647.
261. Pettersson ML, Nedstrand E, Bladh M, Svanberg AS, Lampic C, Sydsjö G. Mothers who have given birth at an advanced age-health status before and after childbirth. Scientific reports. 2020;10(1):9739.
262. Daoud N, O'Campo P, Anderson K, Agbaria AK, Shoham-Vardi I. The social ecology of maternal infant care in socially and economically marginalized community in southern Israel. Health Education Research. 2012;27(6):1018-30.
263. Khan MN, Harris ML, Shifti DM, Laar AS, Loxton D. Effects of unintended pregnancy on maternal healthcare services utilization in low-and lower-middle-income countries: systematic review and meta-analysis. International journal of public health. 2019;64:743-54.

264. Aksünger N, De Sanctis T, Waiyaiya E, van Doeveren R, van der Graaf M, Janssens W. What prevents pregnant women from adhering to the continuum of maternal care? Evidence on interrelated mechanisms from a cohort study in Kenya. BMJ open. 2022;12(1):e050670.
265. de Graft-Johnson J, Kerber K, Tinker A, Otchere S, Narayanan I, Shoo R. The maternal, newborn, and child health: Continuum of care Opportunities for Africa's newborns: Practical data, policy and programmatic support for newborn care in Africa; 2006. Back to cited text. (14):23-36.
266. Holt-Nguyen C. When to Use Logistic Regression vs. Deep Learning for Business Problems 2023 [Available from: https://medium.com/accelerated-analyst/when-to-use-logistic-regression-vs-deep-learning-for-business-problems-589359d49600.
267. Mooney SJ, Keil AP, Westreich DJ. Thirteen questions about using machine learning in causal research (you won't believe the answer to number 10!). American Journal of Epidemiology. 2021;190(8):1476-82.
268. MacNell N, Feinstein L, Wilkerson J, Salo PM, Molsberry SA, Fessler MB, et al. Implementing machine learning methods with complex survey data: Lessons learned on the impacts of accounting sampling weights in gradient boosting. Plos one. 2023;18(1):e0280387.
269. Kanyangarara M, Munos MK, Walker N. Quality of antenatal care service provision in health facilities across sub–Saharan Africa: Evidence from nationally representative health facility assessments. Journal of Global Health. 2017;7(2).