Towards a new method for evaluating large-scale maternal health programmes: measuring implementation strength of focused antenatal care and emergency obstetric care in Tanzania

Gregory Sondo Kabadi

Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy

University of London
January 2015

Department of Disease Control
Faculty of Infectious Diseases

London School of Hygiene and Tropical Medicine

Funded by the Ifakara Health Institute, Tanzania

Declaration by Candidate

I, Gregory, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: Date:

Full name: Gregory Sondo Kabadi

Abstract

Measuring the strength of public health programmes may reveal whether and how some programmes have an impact on target populations and others do not. Programme implementation strength (also known as programme intensity) refers to quantitative measure reflecting programme inputs, processes, and their duration. Measuring programme strength requires an understanding of how programmes work and involves defining measurable concepts, identifying sources of programme data and close programme follow-up. There are no standardized methods for measuring programme strength.

This thesis developed and tested an approach for estimating programme strength for use in evaluating large-scale maternal health programmes in low- and middle-income countries. It used focused antenatal care (FANC) and emergency obstetric care (EmOC) as tracer programmes, with WHO's health-system-building blocks as programme components. The thesis used mixed methods including: developing a weighting scheme through opinions from maternal health experts, collecting FANC and EmOC data from 23 districts on programme strength, programme coverage, and programme contextual factors, using government official statistics, and using routine data from a central database. The thesis also tested the content and face validity of the approach.

Results from experts showed that, even though all six WHO blocks were required in programme implementation, human resources was given relatively higher weights than the other programme components. While the overall programme strength in districts scored an average of 41% (FANC) and 40% (EmOC), the overall programme coverage scored an average of 80% (FANC) and 64% (EmOC). Contextual factors significantly associated with the programmes included: total fertility rate, female literacy, water, sanitation, and famine. The content and face validity were both rated "very good". This work aims to contribute towards an efficient way of evaluating large-scale maternal health programmes in low- and middle-income countries. The approach could also be of interest especially to district health management authorities for improving health programmes.

Table of Contents

Declaration by Candidate **2**

Abstract **3**

Table of Contents **4**

List of Tables **10**

List of Figures **12**

Abbreviations **13**

Acknowledgements **15**

Part I: Introduction **17**

Chapter 1: Background and Rationale **18**

1.1 Background **18**

1.1.1 Global commitments in maternal health 19

1.1.2 Challenges of maternal health programmes in developing countries 20

1.1.3 The need for evaluating maternal health programmes 23

1.2 Programme evaluation designs **24**

1.3 Programme implementation strength **29**

1.3.1 Dose 29

1.3.2 Duration 30

1.3.3 Specificity 30

1.3.4 Intensity 31

1.3.5 Lack of standard methodologies 31

1.4 The need for composite indices **32**

1.4.1 The Maternal and Neonatal Programme Effort Index 33

1.4.2 Quality of Life Index by the Economist Intelligence Unit 34

1.5 Overview of Focused Antenatal Care **35**

1.6 Overview of Emergency Obstetric Care **37**

1.7 The conceptual framework **41**

1.7.1 Introduction 41

1.7.2 Programme implementation strength 41

1.7.3 Programme results 44

1.7.4 Contextual factors 44

1.8 Research aim and objectives **45**

Part II: Methods **49**

Chapter 2: Methods Overview **50**

2.1 Overview **50**

2.2 The Sentinel Panel of Districts **50**
2.3 Indicators of interest and data sources **53**
2.4 Design of research tools **56**
2.5 Training, logistics and data collection **57**
2.5.1 Research assistants 57
2.5.2 FBIS coordinators 58
2.5.3 SARA survey training 59
2.6 Overview of methods **59**
2.6.1 Introducing the implementation strength method 61
2.6.2 Contribution of implementation strength in evaluations 64
2.6.3 The concept of minimal essential indicators 64
2.6.4 Developing the weighting scheme 66
2.6.5 Unweighted programme data 66
2.6.6 Weighted implementation strength scores 67
2.6.7 FANC and EmOC utilization and coverage 68
2.6.8 Contextual factors 68
2.6.9 Illustrative dose-response analysis 69
2.6.10 Content and face validity 70
2.7 Data management **71**
2.7.1 Primary data 71
2.7.2 Secondary data 72
2.8 Data completeness **72**
2.9 Ethical considerations **72**
Chapter 3: Developing the Weighting Scheme for Implementation Strength **74**
3.1 Introduction **74**
3.2 Methods **75**
3.2.1 Design 75
3.2.2 Sample selection 75
3.3 Experts' tool: design, piloting and administration **77**
3.4 Data analysis **82**
3.4.1 Missing values 82
3.4.2 Experts' preferences 82
3.5 Results **82**
3.5.1 Participants 82
3.5.2 Components' contribution to FANC strength 84
3.5.3 EmOC programme implementation strength 88
3.5.4 Priority ranking of FANC activities 91
3.5.5 Priority ranking of EmOC activities 94

3.6 Discussion **94**
3.7 Limitations **96**
3.8 Conclusion **97**
Part III: Results **98**
Chapter 4: Unweighted Scores of the Implementation Strength **99**
4.1 Overview **99**
4.2 Methods **99**
4.2.1 Design and study area 99
4.2.2 Data collection 100
4.2.3 Data quality assessment and data cleaning 104
4.2.4 Data analysis 106
4.3 Results **108**
4.3.1 Human Resources 108
4.3.2 Drugs and Supplies 111
4.3.3 Service delivery 111
4.3.4 Health financing 113
4.3.5 Health information system 116
4.3.6 Leadership and governance 118
4.4 Discussion **119**
4.5 Conclusion **124**
Chapter 5: Weighted Scores of the Implementation Strength **126**
5.1 Introduction **126**
5.2 Summary steps for constructing a composite index **127**
5.3 Methods **129**
5.3.1 Step 1: Theoretical framework 129
5.3.2 Step 2: Selection of indicators 129
5.3.3 Step 3: Imputing missing data 130
5.3.4 Step 4: Multivariable analysis using PCA 131
5.3.5 Step 5: Normalising data 131
5.3.6 Step 6: Aggregating scores and applying the weights 132
5.3.7 Step 7: Uncertainty and Sensitivity Analysis 134
5.3.8 Step 8: Back to the data (Ranking districts) 135
5.3.9 Step 9: Link to other indicators 135
5.3.10 Step 10: Visualisation of the results 135
5.4 Results **136**
5.4.1 Overall scores and district rankings 136
5.4.2 Principal Component Analysis 144
5.4.3 Sensitivity Analysis 146

5.4.4 Reliability 148
5.5 Discussion 148
5.6 Conclusion 152
Chapter 6: Coverage of FANC and EmOC Programmes 154
6.1 Introduction 154
6.2 Methods 156
6.2.1 Study design 156
6.2.2 Variables of interest and data sources 156
6.2.3 Data quality assessment and cleaning 158
6.2.4 Analytical methods 159
6.3 Results 160
6.3.1 Coverage and utilization of antenatal care 160
6.3.2 Coverage of EmOC services 163
6.4 Discussion 166
6.5 Limitations 168
6.6 Conclusion 170
Chapter 7: Contextual factors relevant to FANC and EmOC 172
7.1 Introduction 172
7.2 Methods 173
7.2.1 Selection of indicators 173
7.2.2 Study design 174
7.2.3 The questionnaire and respondents 174
7.2.4 Other programmes and initiatives 174
7.2.5 Water and sanitation 175
7.2.6 Socioeconomic factors 175
7.2.7 Demographic factors 176
7.2.8 Transport and communication 176
7.2.9 Natural disasters 177
7.2.10 Completeness of data and analytical plan 177
7.3 Results 178
7.3.1 Total fertility, infant mortality, and <5 mortality rates 178
7.3.2 Female literacy levels 180
7.3.3 Water and sanitation 180
7.3.4 Transport and communication 183
7.3.5 Floods, drought, outbreaks, and famine 184
7.3.6 Other non-FANC/EmOC projects/initiatives 186
7.4 Discussion 186
7.5 Limitations 193

7.6 Conclusion **193**
Chapter 8: Illustrative dose-response analysis **195**
8.1 Introduction **195**
8.2 Analytical approach **196**
8.3 Results **197**
8.3.1 Univariable analysis 197
8.3.2 Multivariable analysis 200
8.3.3 Dose-response analysis (composite scores) 201
8.4 Discussion **201**
8.3 Limitations **204**
8.4 Conclusion **205**
Part IV: Validity and General Discussion **206**
Chapter 9: Content and face validity of the approach **207**
9.1 Introduction **207**
9.2 Methods **208**
9.2.1 Recruitment of participants 208
9.2.2 Content and face validity testing 209
9.2.3 Analytical approach 212
9.3 Results **213**
9.4 Discussion **218**
9.5 Conclusion **222**
Chapter 10: Application of the implementation strength to programme evaluation and maternal health policy **223**
10.1 The approach's validity **223**
10.1.1 The approach and its key results 223
10.1.2 A realist evaluation's view of the implementation strength approach 229
10.1.3 A theory of change's view of the implementation strength approach 230
10.1.4 Capturing equity dimensions in the implementation strength approach 231
10.2 Potential applicability **232**
10.3 Overall strength and coverage in districts **233**
10.4 Strengthening FANC and EmOC programmes **236**
10.5 The role of partnerships in implementing programmes **238**
10.6 Bias, limitations and methodological challenges **240**
10.7 Conclusion and the way forward **242**
Appendices **245**
Appendix 1: Maternal Health Experts Tool (Implentation Strength Instrument) **246**

Appendix 2: Data Collection Form for Contextual Factors in a District . 251
Appendix 3: Indicators for Data Extraction from DHIS Database 253
Appendix 4: Assessing Content &Face Validity of the Tool for Measuring Implementation Strength 254
Bibliography 255

List of Tables

Table 1.1: Signal functions for identifying Basic and Comprehensive EmOC services..... **38**
Table 1.2: Emergency Obstetric Care Indicators **40**
Table 2.1: Urban and rural districts of the Sentinel Panel of Districts..... **52**
Table 2.2: Initially proposed FANC and EmOC indicators for monitoring programme results **54**
Table 2.3: Proxy indicators for the implementation strength of FANC and EmOC programmes **55**
Table 2.4: Indicators for FANC and EmOC programme results **55**
Table 2.5: Contextual factors and their corresponding proxy indicators‡ **56**
Table 2.6: Linkage between methods, research objectives and data sources **60**
Table 2.7: Types and number of questionnaires before and after data collection, by objective **73**
Table 3.1: Final list of programme activities/elements for programme components..... **80**
Table 3.2: Experts participation by designation and health system level..... **83**
Table 3.3: Analysis of variance of means, participants' views, FANC programme **87**
Table 3.4: Analysis of variance of the means, participants' views, EmOC programme **90**
Table 3.5: Facility-level means priority rank of activities/elements within FANC components, by levels **92**
Table 3.6: Mean priority rank of activities/elements within EmOC components, by levels **93**
Table 4.1: Programme component indicators on the number of health workers and those specifically trained on FANC and BEmOC..... **109**
Table 4.2: Percentage of surveyed health facilities with available Drugs and Supplies on the day of visit..... **110**
Table 4.3: Implementation strength indicators for the service delivery component **112**
Table 4.4: Overall District Revenues in Tanzanian Shillings for 2011 ..... **114**
Table 4.5: Overall District Expenditure in Thousands of Tanzanian Shillings..... **115**
Table 4.6: Implementation strength indicators for the Health Information System component **117**
Table 4.7: Implementation strength on the Leadership and Governance component..... **118**
Table 5.1: Steps/guidelines for building a composite indicator **128**
Table 5.2: Overall Composite Scores and Ranking of Study Districts on FANC Programme Implementation Strength..... **137**
Table 5.3: Overall Composite Scores and Ranking of Study Districts on EmOC Programme Implementation Strength..... **138**
Table 5.4: FANC Principal Components with eigenvalues >1 ..... **145**
Table 5.5: EmOC Principal Components with eigenvalues >1..... **146**

Table 5.6: Cronbach's alpha scores for FANC programme components, with overall test scale, standardised to mean=0 and variance=1 ..... **148**
Table 5.7: Cronbach's alpha scores for EmOC programme components, with overall test scale, standardised to mean=0 and variance=1 ..... **148**
Table 6.1: Coverage of antenatal care and utilization of FANC services in study districts for 2010 and 2011 ..... **161**
Table 6.2: EmOC service coverage in study districts (with 2010 and 2011 routine data and 2012 survey signal functions) **165**
Table 7.1: Demographic factors on TFR, IMR, and U5MR for 2010 and 2011 ..... **179**
Table 7.2: Socioeconomic factors on proportion of literate females in study districts for 2010 and 2011 ..... **181**
Table 7.3: Water and sanitation factors: Proportion of households using drinking water from improved water sources and improved toilet facilities in 2010 and 2011 ..... **182**
Table 7.4: Infrastructural factors on access within districts to transport and communication networks for 2010 and 2011 ..... **184**
Table 7.5: Percentage of households affected by natural disasters in 2010 and 2011 ..... **185**
Table 7.6: Public health system factors based on other projects and initiatives operating in study districts in 2010 and 2011 ..... **187**
Table 8.1: Univariable analysis for associations between the Overall FANC Coverage and the FANC Components & Contextual Variables in 23 districts **198**
Table 8.2: Univariable analysis for associations between the Overall EmOC Coverage and the EmOC Components & Contextual Variables in 23 districts **199**
Table 8.3: Multivariable analysis between overall programme coverage scores and associated predictor variables **200**
Table 9.1: Number expected and level of participants for the assessment survey **210**
Table 9.2: Content and Face validity parameters assessed **211**
Table 9.3: Demographic characteristics of participants (n=42) **213**
Table 9.4: Results of content and face validity of the implementation strength tool (number of times each Likert item was scored, mean, std dev, min and max) **216**
Table 9.5: Weighted Kappa scores for content validity **217**
Table 9.6: Weighted Kappa scores for face validity **217**
Table 9.7: Weighted Kappa scores for both content and face validity **217**
Table 9.8: Selected Comments from Participants on each component of the assessment tool **219**
Table 10.1: Key evaluation approaches for large scale public health implementation, with pros and cons **227**
Table 10.2: Overall Summary of Results **234**

List of Figures

Figure 1.1: The Conceptual Framework **42**
Figure 1.2: Thesis Structure **48**
Figure 2.1: Geographical location of study districts **52**
Figure 2.2: Stages of the implementation strength approach **62**
Figure 3.1: Modified Delphi process used to identify key elements/activities for each FANC and EmOC programme components **79**
Figure 3.2: Frequency distributions (& normal densities) of FANC components (n=210) .. **84**
Figure 3.3: Box plots of percentage weights of FANC components (n=210) **85**
Figure 3.4: Mean percentage contribution of FANC components to implementation strength, by participant level, with overall means, standard deviations (SD), 95% CI **86**
Figure 3.5: An example of how participants assigned weights to individual components from their experience- or level-based opinions **87**
Figure 3.6: Frequency distributions (& normal densities) of EmOC component weights .. **88**
Figure 3.7: Box plots of percentage weights of EmOC components (n=214) **89**
Figure 3.8: Mean percentage contribution of EmOC components to implementation strength, by participant level, with overall means, standard deviations (SD), 95% CI **89**
Figure 4.1: Ordering and flow of supplies under the Integrated Logistics System **121**
Figure 5.1: Study districts performance on FANC & EmOC HR **140**
Figure 5.2: Study districts performance on FANC & EmOC Financing **140**
Figure 5.3: Study districts performance on FANC Service Delivery **141**
Figure 5.4: Study districts performance on EmOC Service Delivery **141**
Figure 5.5: Study districts performance on FANC Drugs & Supplies **142**
Figure 5.6: Study districts performance on EmOC Drugs & Supplies **142**
Figure 5.7: Performance on FANC & EmOC Health Info System **143**
Figure 5.8: Performance on FANC & EmOC Leadership & Governance **143**
Figure 5.9: Sensitivity analyses for the association between the overall FANC implementation strength and programme components **147**
Figure 5.10: Sensitivity analyses for the association between the overall EmOC implementation strength and programme components **147**
Figure 7.1: Dar es Salaam's infrastructural and residential devastation by floods **185**
Figure 8.1: Dose-response between FANC coverage and FANC strength **202**
Figure 8.2: Dose-response between EmOC coverage and EmOC strength **202**
Figure 9.1: Graphs showing 7 Likert scale percentage scores of content validity items (numbers 1 through 4) on algorithm, content and instructions **215**
Figure 9.2: Graphs showing the 7 Likert scale percentage scores of content validity items (number 5 and 6) and face validity items (number 7 and 8) **215**

Abbreviations

AMDD	Averting Maternal Death and Disability
AMREF	African Medical and Research Fund
ANC	Antenatal Care
CCHP	Council's Comprehensive Health Plan
CHMT	Council Health Management Team
DHIS	District Health Information System
DMO	District Medical Officer
EmOC	Emergency Obstetric Care
BEmOC	Basic Emergency Obstetric Care
CEmOC	Comprehensive Emergency Obstetric Care
FANC	Focused Antenatal Care
FBIS	Facility-based Information System
HIV/AIDS	Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome
HMIS	Health Management Information System
IHI	Ifakara Health Institute
IRB	Institutional Review Board
LSHTM	London School of Hygiene and Tropical Medicine
M&E	Monitoring and Evaluation
MDG	Millennium Development Goals
MMR	Maternal Mortality Ratio
MNH	Maternal and Newborn Health
MOHSW	Mainland Tanzania's Ministry of Health and Social Welfare
NEP	National Evaluation Platform
NIMR	Tanzania's National Institute for Medical Research
PCA	Principal Component Analysis
PEPFAR	United States President's Emergency Plan for AIDS Relief
PMTCT	Prevention of Mother-To-Child Transmission of HIV
RCH	Reproductive and Child Health
RCH-Co	Reproductive and Child Health Coordinator
RCHS	Reproductive and Child Health Section of the MOHSW

SARA	Service Availability and Readiness Assessment
SAVVY	Sample Vital Registration with Verbal Autopsy
SPD	Sentinel Panel of Districts
TDHS	Tanzania's Demographic Household Survey
UNFPA	United Nations Population Fund
UNICEF	United Nations Children Fund
WHO	The World Health Organization

Acknowledgements

I would like to thank my supervisor, Prof. Joanna Schellenberg for her challenges, guidance, support, and encouragement throughout my Ph.D. Thank you so much Joanna – "I have found my slippers at last"! I am grateful to my advisory panel members, Prof. Oona Campbell, Dr. Josephine Borghi, and Mr. Jim Todd for their useful comments. I am also grateful to Dr Eveline Geubbels and Dr Mwidimi Ndosi for their occasional advice on my thesis as well as to Ms Jane Bruce for her last minute analysis assistance on weighted kappa statistics. Special thanks should go to Dr. Honorati Masanja for recruiting me from Jhpiego to help him with the SAVVY project and eventually agreeing to let me do my PhD while he still needed my support on the ground. I am also indebted to the Ifakara Health Institute for the financial support without which, it would have not been easy to come to this school. Special appreciation goes to the London School of Hygiene and Tropical Medicine for the 'collaborative site' arrangement with the Ifakara Health Institute that subsidised international fees.

I am grateful to my research assistants who helped me with data collection and their energy for travelling around the 23 SPD districts: Mariam Mashombo, Janenah Mkinga, Wulstan Muhonda, Hendrix Haidhuru, Abraham Madenge, Francis Ndabi, Happiness Moshi, Khadija Kawembele, and Martha Swai. Special thanks to the 23 'FBIS' district coordinators for providing the much-needed assistance to my research assistants and for themselves helping with collecting the data for the contextual factors. I am grateful to all District Medical Officers and their CHMT members, who provided assistance in one way or another towards the work of this thesis.

In London, I spent many 'writing camps' away from my family, most of the time at the Smart Russell Square Hostel, 71 Guildford Street, and I would like to thank the hostel staff for chit-chat stories that gave me a day's relief after the long hours of writing . I am also grateful to my fellow Tanzanian and LSHTM students at different times for their company at school and for the many good times: Frank Mng'ong'o, Dr Joel Francis, Ameer Shaheed, and Zawadi Mageni, to mention a few. My special gratitude go to Jackson Kapama, Brian Ngelangela, Pastor Emmanuel Chatawe, Renatus Mgetta and their families for providing me and my family a warm Tanzanian friendship and neighbourhood in the West Midlands – a Dar es Salaam away from Tanzania.

I would like to thank my late parents Mzee Julius Ikombe Kabadi and Mama Sophia Limi Sondo for laying the foundation of my education. Special thanks to my dad who helped me get along with the multiplication tables ahead of other children at school even if he had to incentivise me with a stick. Concerned about my education after primary school, my dad sent me to Mwalimu Kelemu of Ng'wangongo – Kishapu who transferred me to Mwasele Primary School in Shinyanga town, where I was the only pupil selected to join a government secondary school. Special thanks to my precious mum who recently passed away on 31st March 2014. On one occasion in 1979, she had me dressed with her *khanga* only so I should not miss school for having no uniform. I love you both and miss you deeply! I also owe my sister Margareth Duke Kabadi a big thank you for supporting my secondary school education while she was herself a college student, Asante Sana! My many thanks go to my other siblings and relatives for their support in different ways. I am also indebted to my mother in-law, Mama Rogathe Deborah Mlay for her support to my wife and the children during the last months of writing this thesis in Oldbury, West Midlands.

Lastly but not least, I would like to thank my loving wife Yvonne, for having endured my frequent absence from her and the children during the years of study. I understand how deserted you felt in taking care of our young family while being a fulltime employee and later being a student yourself, yet, you endured it all! The journey towards this thesis has witnessed the arrival of my children: Karen was born three days after I registered for PhD. Michael was born during my data collection year, and Ian, was born during the final writing stages. I will always hold at heart these fond memories.

...and to Him, in whom "all things consist"; who, "in Him we live, and move, and have our being"...

Part I: Introduction

Chapter 1: Background and Rationale

"Without these data, we have no reliable way of knowing whether interventions are working, and whether development aid is producing the desired health outcomes. This is part of our job: to be accountable. We cannot be fully accountable without research"
– Dr Margaret Chan, Director General, WHO, 2007

1.1 Background

While opening the Fourth Global Meeting of Heads of World Health Organisation (WHO) Country Offices in Geneva in 2007, the WHO Director-General, Dr Margaret Chan underscored the need for generating evidence-based knowledge on whether health interventions produce desired health outcomes.[1]

According to the WHO fact sheet on maternal mortality: nearly 800 women die every day globally due to preventable pregnancy- and childbirth-related causes, 99% of the maternal mortality take place in developing countries in women who are either poor or live in rural areas.[2] In 2000, 189 countries convened at the United Nations and agreed to commit themselves into eight goals, the Millennium Development Goals (MDG) for achievement by 2015.[3] One of the eight goals, goal number five, is to improve maternal health. To monitor its achievement over the years, MDG 5 has two targets: target 5a is to reduce the maternal mortality ratio by three quarters, between 1990 and 2015 and target 5b, is to achieve universal coverage of reproductive health by 2015. By 2013, maternal mortality was estimated to have dropped by 45% globally.[4] Alarmingly, 62% of this drop in maternal deaths occurred in the sub-Saharan African region. This decline is largely attributed to financial and political commitments by individual countries and various international organisations. Section 1.1.1 examines briefly such commitments directed to maternal health since the start of the MDGs.

1.1.1 Global commitments in maternal health

For at least the last two decades, there has been an unprecedented increase in investment in global health. According to estimates made by the Institute for Health Metrics and Evaluation, global investment in health in the form of development assistance on health (DAH) nearly doubled from US$ 5.7 billion in 1990 to US$ 10.8 billion in 2001, and almost tripled to US$ 28.1 billion by 2012.[5] Most aid on health to low- and middle-income countries (LMICs) comes through DAH. The increase in global health investment has mostly been due to increased donor funding by developed nations as a result of the MDGs and supported by various funders such the Global Fund to fight AIDS, Tuberculosis and Malaria, the Global Alliance on Vaccines and Immunisation, the Bill and Melinda Gates Foundation, and several other private sources of funding including non-governmental organisations and corporations.[6] [7] Investment in maternal health is associated with numerous benefits – the primary of which is to prevent maternal deaths by saving mothers.[8] By saving mothers, investment in maternal health also means increasing chances of saving their newborns.[9] If a newly delivered mother has other older children, they will also benefit by having a healthy mother to tend to their needs.[10] [11] Tackling maternal health issues also has consequential effects on tackling poverty-related issues as healthy mothers can participate in economic activities for their households.[12] In addition, improving maternal health outcomes has a positive reflection of a country's capability on the performance of its health system.[13] [14] [15]

The "Global Investment Framework" has recently reported interesting estimates of economic and social benefits associated with specific investment in women and children's health.[16] The framework reports that seventy-four high-burden countries have over 95% of the global maternal and child mortality burden. The Global Investment Framework also estimated that increasing the 2013 investments in women and children's health by US$ 5 per person per year in the 74 high-burden countries would yield up to nine times the economic and social benefits by 2035. The beneficial returns include increased growth of gross domestic product (by means of improved labour productivity) and prevention of five million maternal deaths, 147 million child deaths, and 32 million stillbirths in the 74 high-burden countries. While the framework's estimates promise a better world for the high-burden countries, the financial commitment by global bodies does not match the anticipated investment in maternal health. For example, the World Bank promised to commit $700m for investment in low- and middle-income countries by the end of 2015,[17]

based on the World Bank's own projections of the developing sub-Saharan African population for 2013 (0.936 bil),[18] this size of commitment is only a fraction ($0.748 per person) compared to the framework's suggested $5 per person – let alone for the entire population of all low– and middle-income countries, and for 2015. It is unclear whether factoring in other financial investment by other global initiatives such as the United States Agency for International Development (USAID), the United Kingdom's department for international development (DFID), the Bill and Melinda Gates Foundation, the Global Fund to fight AIDS, Tuberculosis, and Malaria, and others, is likely to match the Global Investment Framework's proposed per capita investment in women's and children health.

1.1.2 Challenges of maternal health programmes in developing countries

Maternal health is an important area of unfinished work among the health-related MDGs.[19] MDGs directly related to health include MDG 4: reduce child mortality, MDG 5: improve maternal health, and, MDG 6: combat HIV/AIDS, malaria and other diseases. Challenges in implementing maternal health programmes in LMIC include, among others, the shortage in human and financial resources, limited political commitment, and the availability of reliable data.[20]

Despite the increase in global health investment, there are still investment gaps in specific areas of maternal health. Such investment gaps contribute to the continuation or increase in maternal deaths due to preventable causes such as obstetric haemorrhage, unsafe abortion and other complications of pregnancy and childbirth.[21] Most causes can be prevented by equipping health facilities with skilled staff, essential infrastructure and supplies, and by ensuring functional referral systems.[22] Financial access to life-saving procedures such as Caesarean sections and blood transfusion can be an obstacle to women in need of these services.[23] [24] [25] Transportation costs for travelling to health facilities and high out-of-pocket payments have been reported to inhibit a substantial number of women from accessing maternal health care services in health facilities.[26] Research has reported that government-financing efforts in public and private health facilities are associated with increased access to skilled delivery attendance and Caesarean sections.[27] To address health financing issues and scaling up coverage of maternal health services, governments, the private sector and local and international

organisations must work harmoniously in increasing budgets, expanding insurance schemes, and providing incentives to target groups. [6 28]

Regarding the shortage of human resources for health, the Global Health Workforce Alliance has recently reported that there is a 'serious shortage of health workers across the world' making it 'one of the most critical constraints' to achieving the health-related Millennium Development Goals (MDGs).[29] This shortage affects not only the provision of maternal health services but also implementation of programmes for the three global priority diseases of malaria, HIV/AIDS, and tuberculosis. Of the 57 countries with an acute shortage of health workers (doctors, nurses, and midwives) reported by the World Health Report in 2006, 36 (63%) were African countries.[30] In particular, the Tanzanian health workforce was reported to have dropped in absolute numbers from 67,600 in 1994/1995 to 48,500 in 2001/2002 due to structural adjustment programmes – a period at which the population grew by 20%.[31] With longstanding shortages of doctors, nurses and midwives, rural areas experience a relatively higher burden of disease compared to urban areas where there are relatively more health workers.[32] The overall shortage and the imbalance in geographical distribution of the health workforce in LMIC are a major impairment to the delivery of maternal health services. National governments and the private sector can alleviate the situation by increasing the production and quality of the health workforce through teaching and training and in providing motivational packages for those working in more challenging areas.[33 34 35]

Lack of political interest is another hurdle in ensuring maternal health takes precedence among other priorities.[36] National leaders and local government alike need to be champions of the efforts in improving maternal health and reducing maternal mortality. A positive example was reported of five countries of Guatemala, Honduras, India, Indonesia and Nigeria (India and Nigeria each being one of the greatest contributors of maternal mortality). Besides variations among the five countries, the general political attention was mainly influenced by the "international and national safe motherhood promoters"[37 38] with Nigeria having low political drive behind maternal health later having its federal government leaders being interested in, and local governments producing champions of maternal health.[39] Likewise, advances in technology are much needed for providing drugs, clinical supplies and other medical expertise, but without political action, reduction of maternal and newborn deaths is likely to take longer than necessary. However, policy makers face numerous national issues – all requiring priority and financial attention for

solutions. To bring maternal health issues onto the national agenda, studies have shown that efforts from international agencies and having national champions (or "effective political entrepreneurs") are among the key influence required to attract political priority.[38 40] Even as we near the end of the MDGs in 2015, accelerating the achievement of MDG-5 (reducing maternal mortality and achieving universal access to reproductive health) still requires national political championing.

Reliable data on the other hand, is of great importance in contributing to effective implementation of maternal health programmes, especially in LMIC. Reliable data include both data generated from routine systems and those from surveys. Even though the quality of the routine data from health facilities in several low- and middle-income countries has been questioned for being "incomplete, inaccurate or untimely"[41] compared to the quality of data from most national surveys, [42 43 44] routine data still controls the majority share of the information needed for managing health programmes. Health facility managers of both public and private facilities have a major responsibility of ensuring not only the quality but also the quantity of data generated from their facilities is improved. Examples of such improvement have been documented in several studies such as a study in Mali and Senegal, which investigated the quality of routine health information system data.[45] After the study instituted appropriate and efficient mechanisms in referral hospitals, monitoring of maternal and perinatal health was significantly improved enabling easier access to 'high-quality' data for programme assessment. Due to targeted investment in information systems and infrastructure, vertical programmes in some LMIC have been reported to collecting and using better data compared to other programmes.[46] Establishing health information systems and appropriate infrastructures for collecting both health facility and survey data could significantly help those responsible for implementing safe motherhood initiatives to monitor and evaluate progress towards effective coverage in target populations.[47 48]

While global initiatives, programmes, and national governments are called for prioritising investment in maternal health, expenditure on data collection systems should not be neglected in such joint efforts for improving maternal health and reducing maternal mortality. There is evidence showing countries are spending part of their financial resources in data collection systems. For example, in 2003, McKinsey and Company estimated that, countries around the globe had spent nearly US$ 1.25–2 billion each year on health information, out of which, low- and middle-income countries spent about US$ 0.75–1 billion while the rest was spent by global

initiatives.[49] In addition to the financial investment in data collection systems, investment should also be made in other health information system areas such as legislating and reinforcing information policies, training sufficient number of personnel for data collection and synthesis across health system levels, building infrastructures and policies for transmission, data storage and data use, as well as central coordination of the health information management system.[50]

1.1.3 The need for evaluating maternal health programmes

Financial investors in the health sector can have several questions regarding the effectiveness of their investment efforts, such as, is the investment we made getting the results we anticipated, are the programmes implementing the interventions we funded becoming more efficient in delivering them? Besides such questions, not only that much of the investment in health takes years to generate the intended results but also its benefits can cross the borders of the intended targets (population health) thereby making it difficult to attribute the true effect of the investment. The several millennial commitments have called for countries to 'do better' with their investments in population health and development. Such millennial efforts include the debt reduction programmes (Heavily Indebted Poor Countries and Poverty Reduction Strategy Paper (HIPC/PRSP),[51] the U.S. President's Emergency Plan for AIDS Relief (PEPFAR),[52] the 3 by 5 initiative,[53] the Global Fund to fight against AIDS, Tuberculosis, and Malaria,[54] the Rollback Malaria (RBM),[55] Stop TB partnership,[56] and the Safe Motherhood Initiative,[57] to mention a few. Most of the global initiatives such as the Millennium Challenge Account, PEPFAR, the MDGs, and others have had M&E plans being central to the monitoring of their efforts with most echoing the central message of being ignorance of the impact as no longer acceptable. One particular example of associating investment in health with results is the performance-based financing mechanism that requires health providers to receive funding based on their performance.[58] [59]

The World Bank defines the two terms, monitoring and evaluation (M&E) as: monitoring "is a continuing function that aims primarily to provide the management and main stakeholders of an ongoing intervention with early indications of progress, or lack thereof, in the achievement of results" and evaluation "is the systematic and objective assessment of an on-going or completed project, program, or policy, and its design, implementation and results".[60] M&E has been considered as the "necessary

foundation" for improving programmes and potentially fast track the achievement of the MDGs and other global efforts in health.[61] For example, the International Health Partnership and other initiatives (IHP+)'s global compact for achieving the three health-related MDGs propagates the use of M&E as a platform to "monitor the implementation of national health strategies" whose results are more likely to show the effectiveness of aid in the health sector and in helping to monitor processes globally.[62] IHP+ also has a common evaluation strategic framework that operates in agreement with the effectiveness of aid set out by the Paris Declaration.[63]

The main purpose of monitoring of programmes is to enable implementers observe whether the implementation protocols are being adhered to and if there is a need to change them for improved programme implementation, whereas, an evaluation is carried out to establish whether and how programme activities (interventions) result in desired outcomes (effects). Programmes plan and implement activities in order to achieve immediate outputs, intermediate outcomes, and ultimate impact. There are at least four benefits arising from evaluating maternal health programmes. The first of which, is to determine whether and how the programme is having some positive effect in line with the expectations. Second, evaluating programmes can help to prompt appropriate courses of action to correct the ongoing implementation. Third, evaluation of programmes can help in the estimation of the implementation costs. Fourth, results from programme evaluations form a body of knowledge for common learning that is useful in designing future programmes. Evaluating maternal health programmes is therefore essential and is an integral part of all the efforts aimed at improving the health of women and reduction of maternal deaths.

1.2 Programme evaluation designs

There are two 'traditional' evaluation designs for comparing programme effect:[64] those looking at the same population before and after programme implementation (known as adequacy-type designs) and those looking at outcomes before and after programme implementation in an 'intervention' group compared with a 'comparison' group (known as plausibility- or probability-type designs). A further two evaluation designs have been mentioned elsewhere: adopters/non-adopters design[65] and the national evaluation platform design.[66] The literature also reveals several other approaches for use in identifying and monitoring the impact-causing processes and activities in complex settings such as those found in large-scale health programmes.

The approaches include (but are not limited to) realist evaluation, process evaluation, and, the theory-of-change-based approaches – discussed later in this section.

Adequacy designs involve comparing outcomes in a particular population between two distinct periods by measuring the change. Adequacy designs require no control groups, as evaluation is based on a single population. They are generally less expensive compared to other evaluation designs. Plausibility and probability designs, on the other hand, compare populations receiving and those not receiving the intervention. These types of evaluation designs involve a comparison group that is supposed to be similar to the intervention group and not in receipt of the intervention. The difference between plausibility and probability designs is that probability designs randomize participants to those receiving the intervention and those who do not.[64] As group members (intervention and comparison groups) are considered to have no systematic differences before the intervention, randomisation in probability designs provides them equal chances of receiving the intervention and therefore minimises selection bias such that any difference arising between group members following the intervention can be attributed to the intervention rather than other factors. While plausibility designs would control the influence of external factors (confounding factors) by having a comparison group, their lack of randomisation as done in probability designs denies participants the equal chance to receiving the intervention thereby making it more likely to bias the results. Overall, plausibility and probability designs tend to be (but not always) more costly than adequacy designs, but evidence drawn from their results is generally stronger than that of adequacy designs.[67]

Kirkwood et al describe the adopters/non-adopters evaluation design as one of a community-level intervention is evaluated by analysing individual-level data – for those who adopted the intervention (adopters) and those who did not adopt the intervention (non-adopters).[65] The adopters/non-adopters evaluation design has one obvious limitation of selection bias in that, adopters are likely to be different from non-adopters with regard to characteristics other than the uptake of the programme, and these characteristics may) influence the likelihood of the health outcome. Acknowledging this, Kirkwood et al suggest using both community-level intervention/control groups and the individual-level adopters/non-adopters.

Realising the limitations of the traditional evaluation designs in detecting the change associated with the large-scale programmes due to the 'spill over effect' of other programmes operating in target populations, Victora et al proposed the 'national evaluation platform.[66] Victora et al argued that the on-going programme scale-up in most countries, especially in low- and middle-income countries, means it has become hard to have a 'pure' comparison district that is not affected by programmes in some way. The proposed national evaluation platform would include: use of the 'district' as a unit of design and analysis due to its central role in the health system and thus in programme implementation; continuous monitoring of indicators; additional data-gathering before, during and after the period of assessment; use of multiple analytical techniques (such as time series analyses); and conducting both interim and summative evaluation analyses. Victora et al also argue that, by collecting data from multiple sources where programmes are implemented, through the platform approach, it may be possible both to compare districts with and without a programme and to conduct dose-response analyses.

It has been argued elsewhere that evaluation of large-scale public health programmes is difficult to conduct as programmes involve several interventions whose "pathways to impact are complex and subject to effect modification".[68] [69]Habicht et al acknowledge that there are different components responsible for impacts seen in evaluations.[70] While randomized controlled trials (thought to be gold standards by some) are designed to have maximum internal validity through controlled environments, they tend to lose external validity because the two groups in a trial (intervention group and control group) are often not typical of the population to which interventions will be applied.[64 68] To increase external validity of evaluations, a detailed knowledge of the local context is required.[71] It is therefore unlikely that results of RCTs will be relevant to all settings.[72] As a result, applying probability designs alone may not provide the information about the exact programme elements producing the impact. Because researchers may avoid using RCTs due to real-life conditions for example political reasons (political leaders preferring certain areas in which to give the intervention) or ethical reasons (benefits are already known and therefore everyone should have immediate access) or practical reasons (choosing areas for ease of logistics),[65] the national evaluation platform has some advantages in evaluating large-scale programmes. While the uncontrolled environment within districts threatens the internal validity of this design, the coverage scope enables it to maximise the external validity through its

use of several analytical techniques (such as time-series regression analysis) in dealing with various data gaps and biases.

Other evaluation methods other those discussed earlier include (but not limited to) process evaluation, realist evaluation, and theory-of-change evaluation. The term "process evaluation" includes a wide variety of approaches. Broadly speaking, in a process evaluation, an evaluator assesses how programme activities are implemented, the types of activities, levels of programme participation, and the quality of implemented programme.[73] To conduct a process evaluation, an evaluator will first determine the type of information required and questions to be answered, followed by a careful selection of the appropriate methods – which can include conducting focus group discussions, structured interviews, surveys, review of records and documentation.[74] Generally, process evaluation is considered a useful tool in measuring programme fidelity (the extent to which a programme is implemented as originally intended by those who developed it).[75] It also contributes to understanding relationships between activities and outcomes. Process evaluation can also be useful in assessing accountability and programme improvement and replication. Depending on the method of inquiry used, process evaluation can be implemented even with limited resource. Due to its focus on evaluating processes and activities, process evaluation is complementary to evaluation of the impact of a programme.

Realist evaluation is a theory-driven approach, which seeks to determine 'what works, for whom, and in what circumstances'. [76] The originators of the this evaluation approach, Pawson and Tilley, were not satisfied with the quasi-experimental/experimental designs of evaluation which they criticised for failing to identify the implementation and contextual issues causing programmes to work differently across various contexts. They desired to identify the underlying contextual factors causing the effectiveness of programmes. As a result, they suggested three areas for which programme evaluations should address: the particular programme contexts, the programme's causal mechanisms, and the programme outcomes.[76] To conduct a realist evaluation, Pawson and Tilley proposed that the first step would be to draw out and formalise programme theories to be tested, through searching documents, interviewing programme planners and implementers, reviewing previous evaluation studies and other literature. The second step is to prepare the "CMO framework" through collection of data (both qualitative and quantitative) in order to cross-examine initial hypotheses (preliminary theories) on the **C**ontexts, **M**echanisms and **O**utcomes. This is followed

by conducting a systematic test using the collected datasets to identify the outcome's successes and failures within and across interventions. The fourth step involves assessing and interpreting the analytical results followed by filling out each of the CMO components.

An example of a study using this approach in combination with participatory appraisal was conducted in North West Pakistan to facilitate developing a nutrition intervention programme with local context.[77] Through a series of focus group discussions involving local health workers, the Context–Mechanism–Outcome framework was developed. Further focus group discussions revealed positive results of the programme including the local population gaining knowledge and spreading the knowledge to others. While the realist evaluation's CMO framework can establish the underlying theories for improved programme outcomes, the evaluation approach cannot estimate the public health impact of a programme using quantitative measures. Therefore, for a balanced evaluation of programmes, and whenever possible, it is important to use a realist evaluation perspective in addition to, rather than as an alternative to, impact evaluation.

In theory-of-change-based approaches, evaluators develop a theory of how the intervention is likely to bring about the expected change. Clark and Taplin allude to the process of developing a theory of change as being "a rigorous yet participatory process whereby groups and stakeholders in a planning process articulate their long-term goals and identify the conditions they believe have to unfold for those goals to be met".[78] In order to prepare a theory of change, Anderson shares five stages: [79] identifying the long-term goal of the project/programme, conducting a "backwards mapping" to identify the preconditions necessary to achieve that goal followed by identifying the interventions that the initiative will perform to create these preconditions. The fourth step is to develop the indicators for each precondition for use in to assessing the performance of the interventions, and, lastly, to write a narrative that summarises the challenges in the theory. By preparing the theory of change, evaluators will improve the design and evaluation process especially of those with complex and multi-faceted interventions. While the theory of change approach to evaluation is essential in planning programmes through clear definitions of change-steps, it also cannot estimate the impact of a programme through quantitative measures and should likewise be used in addition to, rather than as being an alternative to, impact evaluation. Several international donor agencies use a theory of change in designing and evaluating the development programmes they

support such as the Bill and Melinda Gates Foundation, the UK Department for International Development and the Comic Relief.[80] [81]

1.3 Programme implementation strength

Implementation strength is a relatively new concept and has been defined as "the quantity of a programme strategy that is carried out at the field/population level and incorporates some elements commonly considered as part of the quality of service delivery".[82] The equivalent of implementation strength in clinical medicine is perhaps that provided by Corday and Pion who defined treatment strength as "the intensity of the intervention relative to the magnitude of the problem that it is intended to correct".[83] Implementation strength is synonymously referred to as programme intensity and differs from implementation fidelity in that programme fidelity is used to evaluate whether a programme was implemented as originally intended.[75] There are few published findings on the strength of interventions and programmes. For example, Miller et al studied the implementation strength and quality of care for the integrated community case management (CCM) of childhood illness intervention in Ethiopia.[84] The study compared intervention and comparison areas. In the intervention area, health workers were trained and were later visited at least once for clinical reinforcement while health posts were supported in purchasing and supplying of drugs and other commodities. Another study in Malawi reported methods of estimating implementation strength for a similar CCM intervention.[85] The study used mobile phone interviews with health surveillance assistants as a validation study to determine whether interviews were a reliable approach for measuring the implementation strength.

Scott and Sechrest pointed out the four dimensions of dose, duration, specificity, and intensity as key dimensions on which to measure the strength of interventions.[86] The following sub-sections briefly discuss the context and methodology of each of these dimensions:

1.3.1 Dose

The dose of a treatment as applied in medical treatments is specified by a doctor as the amount of the drug administered to a patient. In the field of psychology, Yeaton and Sechrest relate a treatment dose to verbal punishment or an electric shock for a misbehaving pupil.[87] A dose of each form of treatment can be termed as weak or

strong depending on its intensity and it length of effect. In public health programmes, a dose can relate to the number of activities implemented or the amount of programme elements received by the intended population. For example, in a maternal health programme for sensitising women on the importance of regular clinic attendance, the number of radio messages, community mobilisation meetings, comic books, and one-on-one motivator meetings can be regarded as programme doses. Differing numbers of doses given to programme recipients can result in different levels of programme effect. A dose-response can generally describe the change in programme effect caused by changing levels of doses and can be used to produce dose-response curves. For example, a study followed up a cohort of over 44,000 US men for two years in order to assess their risk of coronary heart disease in relation to the amount, type, and intensity of their physical activities.[88] Risk ratios corresponding to the quintiles of metabolic equivalent tasks for total physical activity adjusted for age, smoking, and other cardiovascular risk factors were 1.0, 0.90, 0.87, 0.83, and 0.70 – which showed that there was a dose-response relationship between the risk of coronary heart disease and physical activities.

1.3.2 Duration

As applied in medicine, treatments can be administered to patients for differing lengths of time. Again, this depends on several factors as judged by the administering physician such as weight, age, or severity of the disease. For example, a malaria dose for an infant will take into account the age and weight of the baby, and may be administered for a shorter period than for an adult. In public health programmes, some health promotion campaigns can take months or years to bring about the effect of the intended message for programme recipients.

1.3.3 Specificity[a]

In order to treat a condition in medicine, it is necessary to understand the appropriate treatment. In public health, to evaluate the strength of a programme, it is important to know the specific programme activities that were implemented to address the objectives of that programme. For example, some of the examples of

---

[a] Specificity mentioned here is not related to the concept of specificity as a measure of validity.

specific activities associated with a maternal health programme providing emergency obstetric care services in health facilities could be monitoring stocks of the essential drugs for preventing sepsis and treating postpartum haemorrhage.[89] [90] [91]

1.3.4 Intensity

Lyman defines dose intensity for chemotherapy as the "unit dose of chemotherapy administered per unit time".[92] Intensity generally measures the amount of drug administered with reference to time. In medicine, depending on the medication, drugs can be administered all at once or they can be taken over a period of days. Scott and Sechrest illustrate intensity in psychology using the frequency of piano lessons a particular child takes – which can be every few days for some or weekly for others.[86] In public health programmes, there is likely an optimum level at which to train midwives on life-saving skills: for example, 10 sessions over a period of 5 days is more intense than same 10 sessions delivered over 4 weeks.

1.3.5 Lack of standard methodologies

There are no standard methodologies for measuring implementation strength, as most approaches are tailored to individual programmes.[93] However, a recent review of the literature found that most approaches to measuring programme strength shared three key steps: clear definitions of components; detailed tool-development procedures; and transparent weighting and scoring systems.[94] To estimate the strength of programme implementation, two general approaches have been applied: 1) using expert judgment and 2) quantifying specific programme elements.[86] For example, McGrew et al[95] developed an index using a multistep approach that involved compiling a list of key programme elements. By using expert judgments, weights of the relative importance of each element were calculated and used to determine ideal doses of each programme element. Commenting on evaluation of family planning intervention, Ross and Stover point out that using local experts "might exaggerate the strengths of a program"[96] and that international experts (outside the country in which programme evaluation is being conducted) "might be influenced by their knowledge of contraceptives and fertility trends and give lower ratings to programmes that they perceive have performed worse".[97]

Ross and Stover give examples of some attempts to measure and report the strength of public health programmes and interventions. Using questionnaires, Ross and Stover asked a team of expert observers from 89 low- and middle-income countries to score programme components of large-scale family health programmes – after which, they summed individual scores into a single "program effort index" to represent overall programme strength.[97] Again, in 2002 and in continuation of methods employed by Ross and Stover, Bulatao and Ross assessed the strength of maternal and neonatal health services in 49 low- and middle-income countries using a composite index they called "maternal and neonatal program effort index".[98] They used between 10 and 25 experts from each country to rate services on a scale of 1% to 100%.

The concept of measuring the strength of programmes is in many ways a complex one and to some extent requires care in designing programme evaluations. Nonetheless, it is particularly important to understand how to develop measures of programme strength in order to advance the knowledge to improve programme delivery and effectiveness. This kind of knowledge is especially needed at this time and era when funding of public health programmes is increasingly requiring evidence from evaluations that there is value for money in investments to tackle public health problems. Lacking this knowledge can only perpetuate excuses for not improving how to deliver public health programmes in effective and efficient ways.

1.4 The need for composite indices

In estimating the strength of programmes, McGrew et al, Ross and Stover, and Bulatao and Ross (in examples above) used what are known as composite indices. A composite index is a variable resulting from a standard combination of other variables.[99] Jacobs et al define a composite indicator as "an aggregated index comprising individual performance indicators. It is an index of relative attainment since it reflects the relative values of the underlying individual performance indicators".[100] Variables used in generating the composite index need to be measuring a common trait, or should be closely related to one another. For example, a composite index measuring nutritional status of a child might involve such variables as frequency of feeding, quality, and type of food and so forth.[101] There are various ways of generating a composite score of which the most popular and standardised methods are those that use unit weighting and those that use factor

and regression analyses.[102] [103] In regression analysis for example, a composite score can be estimated by setting the regression model with the dependent variable for which to estimate, its identified subcomponents as independent variables and the corresponding coefficients as weights.[104]

The Organisation for Economic Co-operation and Development (OECD) handbook on constructing composite indicators suggests ten steps for constructing a composite index:[105] building it on the basis of theoretical framework, selecting the variables, imputing missing data, conducting multivariate analysis, normalising the data, weighting and aggregating, conducting sensitivity analysis, decomposing the composite indicator to help extend analysis, linking the composite indicator to other variables, and presenting and disseminating results. By summarising a number of variables into a single score, composite indices help to portray the 'big picture' better than would their constituent variables.[106] Based on its articulated procedures, this study uses the ten OECD steps in preparing composite indices in Chapter 5. Due to the many variables involved in estimating the strength of public health programmes, use of composite indices is appealing because it gives a standardised and objective way of producing robust summaries. More discussion on composite indicators is presented in Chapter 5 including how OECD steps were followed to generate composite scores of FANC and EmOC programmes (Section 5.1), as well as on the advantages and disadvantages of composite indicators (Section 5.5). Below are two examples of composite indices relating to health and social life briefly discussed as a simple illustration of how composite indices are constructed and how they involve a number of variables:

1.4.1 The Maternal and Neonatal Programme Effort Index

The Maternal and Neonatal Program Effort Index (MNPI) study was first implemented in 49 countries by Measure Evaluation.[107] It used experts to rate particular areas of maternal and neonatal health. The study used a questionnaire with 13 components, incorporating 81 items. The 13 components were:

- capacities of health centres;
- capacities of district hospitals;
- percent of women with access;
- care at antenatal visits;
- care at delivery;
- care for newborns;
- family planning at health centres;
- family planning at district hospitals;

- policies toward safe pregnancy;
- resources;
- information and education;
- training;
- monitoring and evaluation.

Respondents (the experts) scored each item with a '0' if it was absent or extremely weak and with a '5' if it was optimal, and with a '1', '2', '3', or '4' for intermediate levels. For ease of comparison, scores were multiplied by 20 to range between 0 and 100. Each component had its item scores averaged to obtain a component score. The composite score was the unweighted mean of all 13-component means. To get the overall weighted measure of impact, each country had its composite score weighted by population. Countries were thus ranked according to respective scores. The MNPI study has since has been used to collectively promote overall performance based on composite scores and has involved more low- and middle-income countries. It has generated debate and may have prompted policy changes in some of the participating countries.[107]

1.4.2 Quality of Life Index by the Economist Intelligence Unit

The quality of life index by the Economist Intelligence Unit is specifically discussed here due to its wide use and its diverging approach in calculating composite scores and its weighting scheme. The quality of life index comprises nine factors that have been used to rank over 100 countries from 2005.[108] In 1999, the Economist Intelligence Unit drew results of 'comparable life-satisfaction' surveys from 74 countries to derive weights for its quality of life index. Using multivariate regression, resulting scores were related to nine quality of life factors that were associated with the surveys. The values of the scores thus represent each country's quality of life index. The nine factors and their associated indicators are:

- Material wellbeing (GDP per person, at power purchasing parity in USD),
- Health (life expectancy at birth, in years),
- Political stability and security (political stability and security ratings),
- Family life (divorce rate (per 1,000 population), converted into index of 1 (lowest divorce rates) to 5 (highest)),
- Community life (dummy variable taking value 1 if country has either high rate of church attendance or trade-union membership; zero otherwise),
- Climate and geography (latitude, to distinguish between warmer and colder climes),

- Job security (unemployment rate, in %),
- Political freedom (average of indices of political and civil liberties, with scale of 1 (completely free) to 7 (unfree),
- Gender equality (ratio of average male and female earnings).

In 2005, the Economist Intelligence Unit produced the quality of life index using the coefficients of the resulting multivariate regression equation. The coefficients of the equation are: the constant (2.7959), GDP per person (0.00003), Life expectancy (0.0448), Political freedom (-0.1052), Job security (-0.0217), Family life (-0.1878), Climate and geography (-1.3534), Political stability (0.1519), Gender equality (0.7423), and Community life (0.3865). Even though the quality of life index generates its results from subjective surveys, its linkage with objective determinants of the quality of life adds credibility to its methods.

There is a general indication from the two illustrated examples that the use of composite indices is a common practice in generating a metric that integrates multiple indicators and that the process of generating a composite score seems to be multifaceted. From a data-driven perspective, the process appears to be at least valid due to its apparent incorporation of theoretical perspectives involving use of relevant indicators as well as engaging appropriate mathematical/statistical techniques. Psychometrically, composite indices would be measuring countries' relevant strengths and weaknesses by ranking them in league tables (not shown here but listed in cited papers). A further discussion of the advantages and disadvantages of composite indicators compared to other indicators is presented in Chapter 5.

1.5 Overview of Focused Antenatal Care

In 2001, a multicentre randomised trial was conducted in Argentina, Cuba, Saudi Arabia, and Thailand to compare "the standard model of antenatal care with a new model that emphasizes actions known to be effective in improving maternal or neonatal outcomes and has fewer clinic visits".[109 110] The general results of the trial found no effect of the new approach compared to the standard approach. However, the new approach was said to be more cost-effective and that when implemented, it would be without 'major resistance from women and providers'. Following these results, the World Health Organization issued a manual on how to provide service using the new model of antenatal care (ANC), now commonly known as the goal-

oriented or focused ANC (FANC) and was considered more applicable in low- and middle-income countries.[111] The first sub-Saharan African countries to adopt WHO-approved FANC were Kenya, Tanzania, Uganda, South Africa, Ghana, Malawi, Zambia and Zimbabwe.[112] Other sub-Saharan African countries have since been scaling up FANC. In 2002, the Tanzanian government issued a policy on provision of antenatal care using the FANC approach.[113]

Unlike the standard ANC that centres care on risk assessment and on more and frequent ANC visits, FANC emphasizes quality over quantity in care of pregnant women in helping to maintain normal pregnancy, prevent complications and facilitate early detection and treatment of complications with a maximum of four ANC visits.[114] For example, in the Tanzanian FANC operational guideline, antenatal care service is organised in five thematic components of history taking, physical examination, laboratory examinations, drug administration and immunization and health education.[113] [115] History taking seeks to identify the woman's individual situation in order to detect evidence of chronic conditions and other problems with potential to cause pregnancy complications. Specialized attention is provided to a pregnant woman with signs of complications who is likely to have more visits to the clinic for close monitoring. During the course of the four visits, pregnant women are provided with preventive services such as immunization against tetanus, prevention of iron deficiency, protection against malaria for women living in malaria endemic countries, presumptive treatment for hookworm, screening, and treatment of syphilis, and prevention of mother to child transmission of HIV (PMTCT). Health education includes health messages and counselling women and families on appropriate pregnancy care and birth preparedness.

A Cochrane review of studies compared the effects of antenatal care packages between the standard approach and the 'FANC' approach for low-risk women on primary outcomes of death of the baby, on several secondary outcomes and on costs to health services.[116] Results showed no evidence of a difference between the two approaches for most of the outcomes assessed except that there was a 14% increase in perinatal mortality for studies using the FANC approach compared to those using the standard approach (95% CI 0% to 31%). Similarly, as was found in the multicentre randomised trial, the review also concluded that the FANC approach was more cost-effective than the standard group, attributing the cause to the smaller number of visits. In Tanzania, several studies have reported on results of a particular aspect of FANC services – most of which were on the quality. For example, the poor

quality of FANC services was reported to be due to poor implementation of the national FANC guidelines, absenteeism of health providers, shortages of FANC-trained staff and ANC supplies, and poor provision of counselling to clients including spending far less time than recommended.[117 118 119 120 121] One study reported that women in different ethnic groups preferred to deliver at home mainly due to poor communication by health providers to women on the importance of delivering in health facilities, but their husbands' decisions were reported to also influence choice of places of delivery.[122]

1.6 Overview of Emergency Obstetric Care

According to the World Health Organization, emergency obstetric care is defined as "the elements of obstetric care needed for the management of normal and complicated pregnancy, delivery and the postpartum period".[123] Pregnant or recently delivered women are subjects of various complications that occur during antenatal, during and after childbirth. Globally, about 300,000 women die due to complications of pregnancy and childbirth each year.[124] In addition, for each woman dying due to pregnancy and childbirth complications, about 20 more women are involved with birth-related disabilities or illnesses such as obstetric fistula, uterine prolapse, anaemia, or infertility.[125] Haemorrhage, sepsis, unsafe abortion, hypertensive disorders, and obstructed labour are the five most occurring complications that cause maternal mortality. Most obstetric complications are either unpredictable or not preventable from occurring, as they are likely to affect even healthy pregnant women.[126] However, most of the complications can be avoided by having all pregnant women attending antenatal care visits for continuous assessment of danger signs and having all deliveries attended by skilled staff who can timely recognize and treat them appropriately.[127]

In order to reduce maternal mortality, emergency obstetric care should therefore be available to all women who develop complications, and that treating obstetric complications should be conducted in health facilities that have sufficient equipment and skilled human resources. Emergency obstetric care services (also widely known as signal functions, Table 1.1) include: parenteral administration of antibiotics, administration of uterotonic drugs such oxytocin, parenteral administration of anticonvulsants for pre-eclampsia and eclampsia, manual removal of the placenta, removal of retained products, assisted vaginal delivery, basic neonatal resuscitation,

blood transfusion and delivery by Caesarean section.[128] Of the nine signal functions of emergency obstetric care, the first seven are classified as basic emergency obstetric care (BEmOC), and comprehensive emergency obstetric care (CEmOC) embodies all BEmOC services plus the other two signal functions of blood transfusion and delivery by Caesarean section. The World Health Organization recommends that lower level facilities (dispensaries and health centres) be capable of offering all seven elements of basic emergency obstetric care and that, hospitals and some health centres with functioning operating theatres be capable of transfusing blood and performing Caesarean section deliveries.

Table 1.1: Signal functions for identifying Basic and Comprehensive EmOC services

No	Basic Services	Comprehensive Services
1	Administer parenteral antibiotics	All basic signal functions (1–7), plus: **8** Perform surgery (Caesarean section) **9** Perform blood transfusion
2	Administer uterotonic drugs (i.e. parenteral oxytocin)	
3	Administer parenteral anticonvulsants for pre-eclampsia and eclampsia (i.e. Magnesium Sulphate).	
4	Manually remove the placenta	
5	Remove retained products (e.g. manual vacuum extraction, dilation and curettage)	
6	Perform assisted vaginal delivery (e.g. vacuum extraction, forceps delivery)	
7	Perform basic neonatal resuscitation (e.g. with bag and mask)	

Source: WHO/UNFPA/UNICEF/AMDD Handbook on monitoring obstetric care[128]

There are eight indicators for monitoring provision and uptake of emergency obstetric care services in populations. The indicators were provided in the World Health Organisation (WHO)/United Nations Population Fund (UNFPA)/United Nations Children's Fund (UNICEF)/ Averting Maternal Death and Disability Program (AMDD) handbook on monitoring emergency obstetric care ("Monitoring emergency obstetric care: a handbook"),[128] which are shown in Table 1.2. Six of these indicators were in the original UNICEF/WHO/UNFPA guideline produced in 1997 with the additional two recently added following new evidence from studies.[129] Table 1.2 also shows descriptions for the indicators, the numerators, denominators and their recommended minimum levels (acceptable level). These indicators are useful in informing whether the target population is actually receiving good quality EmOC services equitably.

Receiving EmOC services that are of good quality and that are equitable can be difficulty to reach all women in need of the services. Evidence generated from research involving both sides of provision and use of obstetric care have reported several issues that should be addressed to achieve the desired coverage of the women. Thaddeus and Maine originally grouped such issues in three phases of delays: delayed decision to seek care, delayed arrival at point of care, and delayed reception of adequate care.[130] Other studies have grouped and presented the issues differently. One example of different grouping of the issues was reported by Kongnyuy et al who named them as barriers and categorised them as those due to health policy, infrastructure, equipment and supplies, the referral system, human resources, health information system, financial barriers, quality of EmOC, and health seeking behaviour.[127] Following the three delays pattern, in the first delay women face several obstacles in deciding to seek care including their limited or lack of knowledge of the pregnant-related risk factors and complications, the timing when they are supposed to seek care, and, other social and economic concerns at the household level. In the second delay, women encounter issues around accessing health facilities including distance, and availability and cost of transport. In the third delay, lack of drugs and supplies, insufficient number of skilled providers and functionless referral systems are the key obstacles for women receiving adequate care in health facilities.

Table 1.2: Emergency Obstetric Care Indicators

No	Indicator	Description	Numerator	Denominator	Acceptable level
1&2 *	Availability of EmOC (national or subnational)	Ratio of EmOC facilities to population and geographical distribution of facilities	No. of facilities in area providing basic or comprehensive EmOC	Population of area divided by 500 000	≥ 5 EmOC facilities per 500 000 population
			No. of facilities in area providing comprehensive EmOC	Population of area divided by 500 000	≥ 1 comprehensive facility per 500 000 population
3	Proportion of all births in EmOC facilities	Proportion of all births in population in EmOC facilities	No. of women giving birth in EmOC facilities in specified period	Expected no. of births in area in same period	Recommended level to be set locally
4	Met need for EmOC	Proportion of women with major direct obstetric complications treated at EmOC facilities	No. of women with major direct obstetric complications treated in EmOC facilities in specified period	Expected no. of women with severe direct obstetric complications in area in same period**	100%
5	Caesarean section as a proportion of all births	Proportion of all births in population by Caesarean section in EmOC facilities	No. of Caesarean sections in EmOC facilities in specified period	Expected no. of births in area in same period	5–15%
6	Direct obstetric case fatality rate	Proportion of women with major direct obstetric complications who die in EmOC facilities	No. of maternal deaths due to direct obstetric causes in EmOC facilities in specified period	No. of women treated for direct obstetric complications in EmOC facilities in same period	< 1%
7	Intrapartum and very early neonatal death rate	Proportion of births that result in an intrapartum or a very early neonatal death within the first 24 hours in EmOC facilities	No. of intrapartum deaths (fresh stillbirths; ≥ 2.5 kg) and very early neonatal deaths (≤24 h; ≥ 2.5 kg) in EmOC facilities in specified period	No. of women giving birth in EmOC facilities in same period	"To be decided"
8	Proportion of maternal deaths due to indirect causes	Percentage of all maternal deaths in EmOC facilities due to indirect causes	No. of maternal deaths due to indirect causes in EmOC facilities in specified period	All maternal deaths (from direct and indirect causes) in EmOC facilities in same period	None set

Source: WHO/UNFPA/UNICEF/AMDD Handbook. * Indicators 1 and 2 involve the same calculations, with data on the corresponding regional population and facility instead of aggregated national data. ** Equal to 15% of expected births in the same area and period.

1.7 The conceptual framework

1.7.1 Introduction

Figure 1.1 shows the conceptual framework for this thesis. The framework was informed by Donabedian's framework for assessing care quality (structures, processes, and outcomes)[131] and the International Health Partnership's M&E framework.[63] The framework shows three main parts: part one is on the implementation strength of FANC and EmOC programmes (comprised of inputs and processes), part two is on programme results (comprised of programme outputs and outcomes), and part three is on contextual (or external) factors. Each part includes names of data sources appearing in **coloured fonts** indicating data retrieval methods. Arrows drawn reflect the unidirectional relationship between the parts. The following sub-sections describe each of the three parts of the conceptual framework. Even though not shown in the framework, the concept relied on use of minimum essential set of indicators. In consideration of limited resources, a minimum set of indicators was required for use in collecting data from the 23 study districts. More details on the concept of minimum essential data is provided in Section 2.6.3.

1.7.2 Programme implementation strength

As previously defined, implementation strength of a programme involves quantifying programme strategies, including elements of quality responsible for delivering the programme to target population.[82] For the current study, a literature search of the PubMed database and Google search engine was conducted to identify the 'programme strategies' related to FANC and EmOC programmes and involved identifying unique features of these programmes.[75 132]

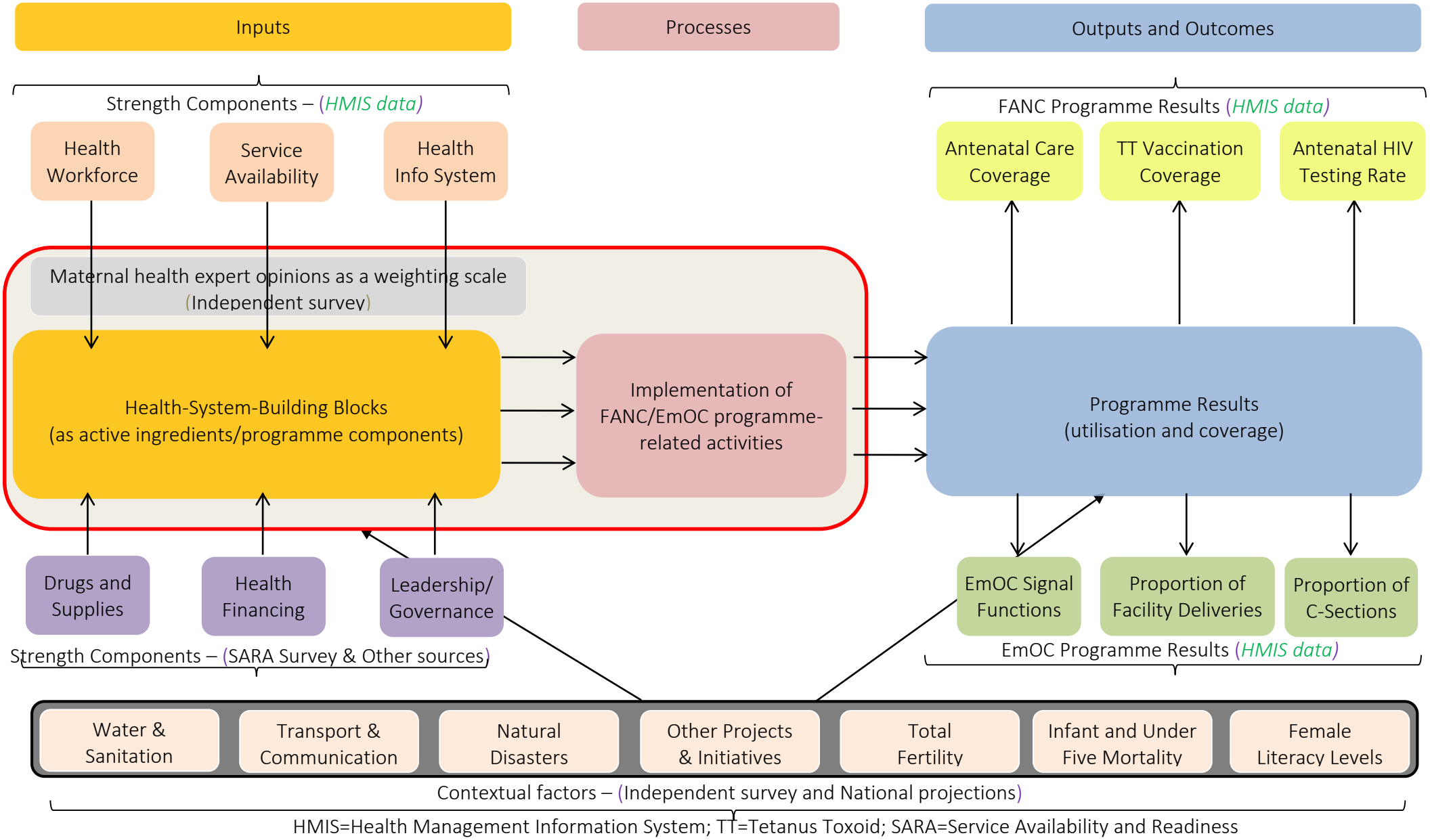


Figure 1.1: The Conceptual Framework

From the review of the literature (whose further explanation appears in Chapter 3), the six WHO health-system-building blocks were eventually judged as being the programme components most appropriately packaged for implementation of the FANC and EmOC programmes. The six building blocks are health workforce, drugs and supplies, service delivery, health financing, health information system and leadership and governance.[133] [134] Having adopted the six building blocks, a further review of the literature (also further explained in Chapter 3) was conducted to identify types of activities or elements constituting each component – an exercise that culminated into developing a survey tool. The survey tool (showing the six building blocks and their corresponding activities/elements) was presented to three senior maternal health experts for an independent review. The experts modified the list by adding and removing some of the activities or elements in components (described further in Chapter 3). A larger survey was conducted involving over 200 maternal health experts who rated the six components and their respective elements – results of which produced the weighting scales (further discussed in Chapter 3).

Processes appearing as part of the implementation strength are programme activities that produce the outputs. For FANC and EmOC programmes, activities can include (but not limited to) training of nurses and midwives, conducting outreach services, ensuring a referral system is functioning, distribution of information, education, and communication materials, conducting morning sessions to pregnant women, distributing ANC and EmOC drugs and supply materials, conducting supportive supervision visits to health facilities, and many others. Details of part one of the conceptual framework (on implementation strength) appear in Chapter 4 and includes the details on how data for the six building blocks were used to estimate the strength of FANC and EmOC programmes. The details also include a discussion of the results.

1.7.3 Programme results

Part two of the conceptual framework is on programme results (outputs and outcomes) that reflect the pattern and scope of access and use of FANC and EmOC services by pregnant and postpartum women as a response to implementation of the programmes. It is important to note that, due to the cross-sectional nature of the research, programme results referred here are only those on utilisation and coverage levels captured at one point in time and do not include those on long-term impact on populations such as reductions in maternal or infant mortality rates. The framework shows the type of the results (in small boxes) and includes the following for FANC programme: the coverage levels of antenatal care in districts, the proportion of antenatal care clients who were tested for HIV/AIDS, and the proportion of antenatal care clients who received between two to five doses of tetanus toxoid (TT) injection. For EmOC, programme results include coverage of the EmOC signal functions, the rate of Caesarean sections performed in a district, and the proportion of births delivered in health facilities. These results items are by no means exhaustive of all utilisation and coverage of FANC and EmOC programmes but a proxy representation of many indicators whose selection is detailed in Chapter 6. Chapter 6 also describes in greater depth the methods used to capture data for the selected indicators, data analysis and the discussion of the results.

1.7.4 Contextual factors

The extent to which a programme is implemented (and likely its strength) is influenced by many factors which can either facilitate the process or impede the implementation process. Such factors are commonly known as contextual factors and are shown in the framework in the underlying box below the other two parts of implementation strength and programme results. In addition, contextual factors can also include factors that influence the outcome but have nothing to do with the programme under study. Due to their likely effect to programme implementation and programme coverage, it is important that data collection also involve the potential contextual factors. The literature gives two reasons why it is important to measure the contextual factors:[135] one, contextual factors can play the role of confounders (for example, they can be associated with the programme components and at the same time be associated with the outcomes), and second, many public

health interventions are context-specific (for example, certain programme activities may be conducted in some part of a district but not all of the target population will be able to access the services due to such factors as natural disasters and transport issues or distance to the health facilities).

A typical example of a programme evaluation that took stock of the contextual factors along with input- and outcome-related factors was the Multi-Country Evaluation of the Integrated Management of Childhood Illness (IMCI) which was conducted in five countries of Bangladesh, Brazil, Peru, Uganda and Tanzania.[136] The evaluation identified two types of contextual factors: those related to the implementation (that is factors related to the health system where IMCI was applied) and those related to the impact measures of the project (including nutrition and mortality levels for children). The study found that the contextual factors played a key role in explaining differences between two of the five case countries (Tanzania and Peru). In addition, documentation of the contextual factors is central to explaining the external validity of programmes – thus describing the extent to which programme results can be generalised in other environments.[68] The current study selected and collected data for a small number of FANC- and EmOC-related contextual factors of different levels and categories as shown in the framework. These include water and sanitation, transport and communication, natural disasters (floods, drought, disease outbreaks, and famine), other projects and initiatives operating in districts, total fertility rate, infant mortality rate, under five mortality rate, and female literacy levels in districts. A further discussion of the contextual factors is presented in Chapter 7 including how the indicators were selected and their data collected, and the analysis and discussion of results.

1.8 Research aim and objectives

The main goal was to develop and test an approach for estimating programme implementation strength that can be used to evaluate large-scale programmes in low- and middle-income countries. The approach used data from two illustrative maternal health programmes of focused antenatal care and emergency obstetric care. The specific objectives were:

1) ***Objective 1:*** To develop an approach for estimating programme implementation strength by:
   - developing a weighting scheme for estimating the strength of focused antenatal care and emergency obstetric care through use of opinions from maternal health experts;
   - and applying the weighting scheme to programme indicators to estimate the implementation strength of the two tracer programmes in a nationally representative sample of 23 study districts.

2) ***Objective 2:*** To assess utilization and coverage of focused antenatal care and emergency obstetric care programmes in the 23 districts using 2-year routine monitoring data (Jan 2010 – Dec 2011).

3) ***Objective 3:*** To investigate and analyse district-level contextual factors as potential confounders of the effect of FANC and EmOC on programme utilization and coverage.

4) ***Objective 4:*** To test the programme implementation strength approach by investigating the association between programme implementation strength and utilization and coverage of the two programmes through an illustrative dose-response analysis.[b]

5) ***Objective 5:*** To test the content and face validity of the programme implementation strength tools.

In order to achieve the objectives, the conceptual framework guides the research of this study in providing answers to the following three essential questions:

1) What are the active components responsible for the implementation of FANC and EmOC programmes in the 23 study districts of Tanzania?

---

[b] Please see Section 2.6.2 on "Illustrative dose-response analysis" for conditions on using multivariable regressions for dose-response associations.

2) What are the likely contextual factors that can confound the implementation and outcomes of FANC and EmOC programmes in the 23 study districts of Tanzania?

3) With cross-sectional data, is there an indication that stronger implementation of FANC and EmOC programmes is associated with higher programme coverage than weaker implementation of the two programmes in the 23 study districts?[c]

The thesis is divided into four parts (Figure 1.2): Part ONE presents the introduction comprising the background and context, a brief overview of the literature, and the thesis objectives (Chapter 1). Part TWO comprises the overview of methods used (Chapter 2) and the detailed description of how the weighting scheme for the two programmes was developed (Chapter 3). Part THREE presents unweighted results of the implementation strength of FANC and EmOC programme data (Chapter 4), results on the composite scores of the weighted programmes data (Chapter 5), results on the coverage of the two programmes in study districts contextual factors (Chapter 6), scores results of the contextual factors (Chapter 7) and the illustration of the dose-response analysis (Chapter 8). Part FOUR includes the content and face validity of the proposed approach (Chapter 9) and concludes with the overall discussion and the implications for policy and future research (Chapter 10).

---

[c] Please see Section 2.6.2 on "Illustrative dose-response analysis" for conditions on using multivariable regressions for dose-response associations.

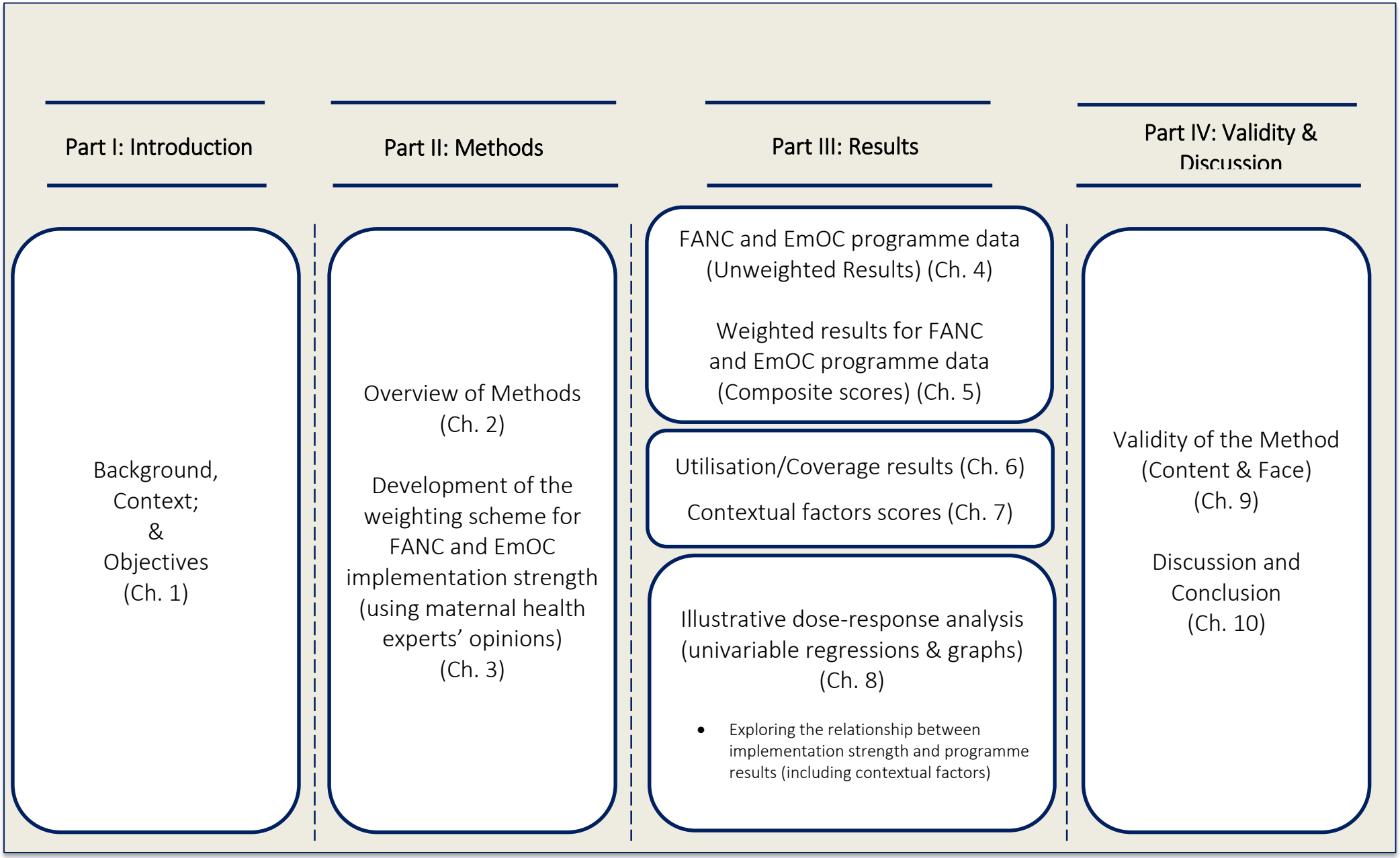


Figure 1.2: Thesis Structure

Part II: Methods

Chapter 2: Methods Overview

2.1 Overview

Chapter 2 provides a brief description of the methods used. Detailed description of the methods for each sub-study is provided in the respective chapters. Section 2.2 briefly introduces the study area and talks about sampling issues. Section 2.3 outlines the indicators and sources of data for programme components and programme results. Design of the research tools is briefly discussed in Section 2.4. Section 2.5 talks about logistics, training, and data collection activities. Section 2.6 presents an overview of the methods employed by each of the sub studies undertaken. Sections 2.7 and 2.8 discuss data management and data completeness issues. Section 2.9 closes the chapter by presenting ethical issues and research funding.

2.2 The Sentinel Panel of Districts

Ifakara Health Institute in Tanzania[d] is implementing the "Sentinel Panel of Districts" which was established in 2009 as a national platform for the evaluation of health programmes and interventions. The sentinel panel of districts consists of two arms: the community-based arm called the "Sample Vital Registration with Verbal Autopsy System" (SAVVY) and the facility-based arm called the "Facility-based Information System" (FBIS). SAVVY is a group of methods that are used for monitoring and reporting of vital events in communities and includes the WHO's verbal autopsy tools for ascertaining probable causes of death in communities.[e] FBIS was set up to collect data from health facilities and aims to improve routine health management information system (HMIS) data collection in health facilities in the

[d] Ifakara Health Institute is an independent, non-profit organisation, registered in Tanzania and led by Tanzanians. The institute conducts a wide range of health-related research, including biomedical and environmental studies, trials of drugs, vaccines and diagnostics, health-systems research, and monitoring and evaluation. More information about the institute can be found at www.ihi.or.tz

[e] More about SAVVY methods can be accessed through Measure Evaluation of the University of North Carolina at http://www.cpc.unc.edu/measure/tools/monitoring-evaluation-systems/savvy

districts for better monitoring of the burden of disease, health outcomes, service delivery, and clinical practice in relation to Tanzania's monitoring and evaluation framework for health sector strategic plans. Much of the data on coverage of FANC and EmOC programmes were drawn from the FBIS arm, discussed later in the results chapter on programmes coverage (Sections 6.2 and 6.3) .

In 2009, the Tanzanian National Bureau of Statistics prepared the sample of districts for the community-based arm, SAVVY. In the National Population and Housing census conducted in 2002, Mainland Tanzania had 21 regions with 119 districts. To date, regions have been further divided, with 4 additional regions and 19 new districts. The sampling frame used by the Bureau of Statistics consisted of the 21 (2002) regions. In preparing the SAVVY sample, the National Bureau of Statistics stratified the regions into eight geographical zones, of which seven were the government's existing administrative zones. The three districts of the City of Dar es Salaam (situated in the East Coast) were treated as regions and the city itself as a special zone due to its high population density. The final sampling frame excluded all special enumeration areas such as schools, hospitals, police and military barracks and seven districts that were involved in demographic surveillance activities. Using its Statistical Master Plan and the 2002 Population and Housing Census dataset, the Bureau of Statistics employed a two-stage probability proportional to size sampling method to obtain a nationally representative sample of 23 districts. Figure 2.1 shows the geographical distribution of the final sampled districts, the Sentinel Panel of Districts (SPD).

SPD districts have a wide range of geographical features including lowlands and highlands. SPD districts were distributed proportionally in each zone based on the number of regions within a zone. Among other variables, districts were stratified by residence: urban/rural (Table 2.1). Overall, SPD districts include a total of about 1,608 dispensaries, 193 health centres, and 87 hospitals – about 35% coverage of Mainland Tanzania's health facilities. Ownership of facilities includes public, private, faith-based, and parastatal organisations.

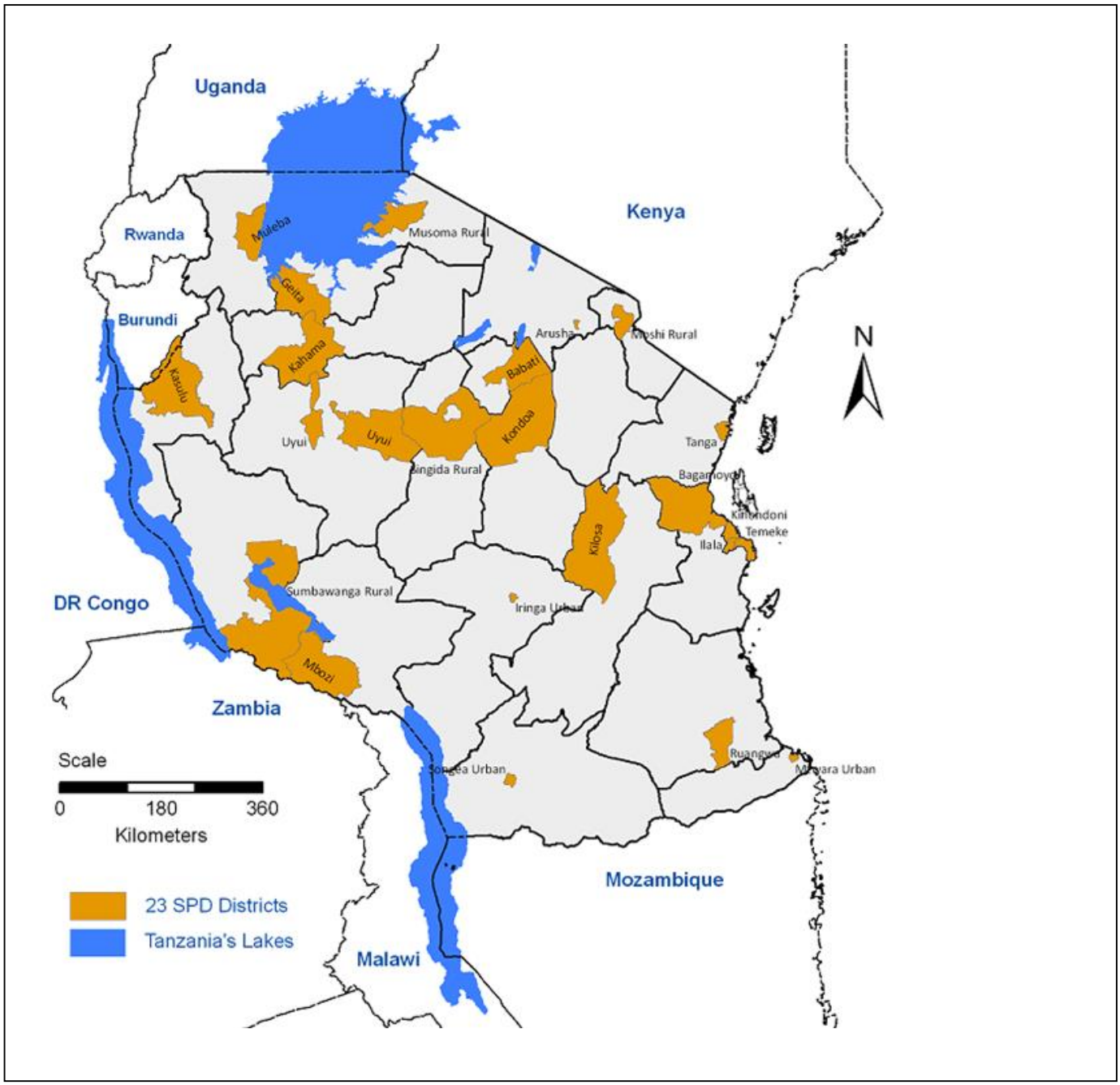


Figure 2.1: Geographical location of study districts

Table 2.1: Urban and rural districts of the Sentinel Panel of Districts

Urban Districts		Rural Districts			
1	Arusha Urban	9	Babati	17	Moshi Rural
2	Ilala	10	Bagamoyo	18	Muleba
3	Iringa Urban	11	Geita	19	Musoma Rural
4	Kinondoni	12	Kahama	20	Ruangwa
5	Mtwara Urban	13	Kasulu	21	Singida Rural
6	Songea Urban	14	Kilosa	22	Sumbawanga Rural
7	Tanga Urban	15	Kondoa	23	Uyui
8	Temeke	16	Mbozi		

2.3 Indicators of interest and data sources

While designing the study and before data collection, an assessment was made of the potential sources of data in study districts for both the implementation strength (the six WHO's blocks, or, programme components) and programme utilisation and coverage (programme results). A careful selection of proxy indicators was made to ensure selected indicators fulfilled the following inclusion criteria: that the indicator would be representative of the group of service (relevance), that the indicator's data sources were credible and data were available, and that data quality was adequate. Table 2.2 shows the list of initially proposed indicators for monitoring the programme results.

Desirable as they were, several indicators could not be included in the final list for having not fulfilled some or all of the inclusion criteria. For example, the proportion of pregnant women receiving 90 or more Iron tablets during current pregnancy was a desirable indicator on the utilization of FANC programme, but after close assessment of the DHIS database, it was found that it would be difficult to get the data from study districts as the indicator had only been introduced by the Ministry of Health and Social Welfare in 2011 and that most facilities in the country had not started collecting relevant data. Tables 2.3, 2.4 and 2.5 show the final list of indicators that satisfied most or all of the inclusion criteria. The Tables also show the indicators' data sources for the implementation strength components, programme results, and contextual factors respectively.

Table 2.2: Initially proposed FANC and EmOC indicators for monitoring programme results

	No	Indicator	Numerator	Denominator		Data Source	Rationale
Antenatal Care	1	Antenatal first visit coverage	ANC first visit	Pregnancies Expected		Reproductive and Child Health Registers for Antenatal Care	The first visit is a "registration" visit for initial procedures to assess/prepare a woman for pregnancy and delivery. This includes full history, examination, initial blood tests and immunisation. Fourth ANC visit shows continuity of care, which is often related to perceived quality. ANC coverage is an indicator of access and use of health care during pregnancy. Syphilis and HIV testing coverage shows quality of care.
	2	Antenatal Fourth visit coverage	ANC fourth visit	Pregnancies Expected			
	3	ANC clients protected from Tetanus	ANC 2-5 tetanus toxoid injection doses	ANC first visit	To estimate population coverage, calculations will also involve a denominator on number of pregnancies		
	4	ANC IPT 2 coverage	ANC IPT 2nd dose given	ANC first visit			
	5	ANC HIV prevalence	ANC HIV tests positive	ANC HIV tests			
	6	ANC Iron supplementation rate	ANC Fe or Fe /+Fol tablets 90 or more given	ANC first visit			
	7	ANC de-worming rate	ANC Mebendazole/Albendazole tablets	ANC first visit			
	8	ANC Anaemia testing rate	ANC haemoglobin test done	ANC first visit			
	9	ANC HIV testing rate	ANC HIV tests done	ANC first visit			
	10	ANC Syphilis testing rate	ANC Syphilis tests done	ANC first visit			
	11	ANC ITN vouchers	ANC recipients of ITN vouchers	ANC first visit			
	12	Stock-out rates of tracer FANC drugs	TT doses, SP, Fe/+Fol tablets, Mebe/Albendazole	Stock-out days per quarter			
Obstetric Care	1	Availability of facilities providing Basic EmOC services	Facilities which have reported all six BEmOC signal functions of complicated delivery within the last 3 months: 1. Parenteral administration of antibiotics 2. Administration of uterotonic drugs 3. Administration of parenteral anticonvulsants for pre-eclampsia and eclampsia 4. Manual removal of the placenta 5. Removal of retained products 6. Assisted vaginal delivery 7. Basic neonatal resuscitation	500,000 Population		Health facility records including Labour and Delivery Ward registers and Maternity home records	Major causes of maternal mortality e.g. Sepsis, pregnancy induced hypertension and (pre) eclampsia need urgent intervention to prevent death Manual removal of placenta stops most postpartum haemorrhage Vacuum extraction can stop delays, prevent ruptured uterus and save foetal lives, etc.
	2	Availability of facilities providing comprehensive EmOC services	Within the last 3 months, facilities that reported all BEmOC functions AND: 8. Caesarean section 9. Blood transfusion	500,000 Population			5-15% of all deliveries require caesarean section. Lower or higher levels require further investigation and action Blood transfusion measures whether the facility has Comprehensive care capacity
	3	Proportion of all births in EmOC facilities	No. of women giving birth in EmOC facilities in specified period	Expected no. of births in area in same period			
	4	Met need for EmOC	No. of women with major direct obstetric Complications treated in EmOC facilities in specified period	Expected no. of women with severe direct obstetric complications in area in same period			Deliveries in the facility is the best way of reducing maternal mortality as deliveries are more likely to be attended by skilled birth attendants
	5	Caesarean section as a proportion of all births	No. of caesarean sections in EmOC facilities in specified period	Expected no. of births in area in same period			
	6	Direct obstetric case fatality rate	No. of maternal deaths due to direct obstetric causes in EmOC facilities in specified period	No. of women treated for direct obstetric complications in EmOC facilities in same period			
	7	Intrapartum and very early neonatal death rate	No. of intrapartum deaths in EmOC facilities in specified period	No. of women giving birth in EmOC facilities in same period			
	8	Proportion of maternal deaths due to indirect causes	No. of maternal deaths due to indirect causes in EmOC facilities in specified period	All maternal deaths (from direct and indirect causes) in EmOC facilities in same period			
	9	Stock-out rates of tracer EmOC drugs	Oxytocin, Ergometrine and Magnesium Sulphate	Stock-out days per quarter			

Table 2.3: Proxy indicators for the implementation strength of FANC and EmOC programmes

No.	Programme Component	Indicator/s	Source of Data
1	Health Workforce	Health worker/population ratio: (doctors, nurses and midwives including non-physician clinicians and lab technicians and pharmacists)	Official statistics from the Ministry of Health and Social Welfare; Population projections from NBS (2010 & 2011)
2	Essential Medicines (Drugs and Supplies)	Availability of ANC tracer drugs: SP drugs, TT injections, Iron and/or Folic acid); Availability of EmOC tracer drugs – Oxytocin, Ergometrine and Magnesium Sulphate)	DHIS/SARA survey
3	Service Delivery	Number and distribution of health facilities per 10,000 population, the number and distribution of inpatient beds per 10,000 population, and number of first antenatal care visits per 10,000 population.	Households estimates from NBS; Number of health facilities in districts from DHIS database/SARA survey
4	Health Information System	Quarterly HMIS reporting rate – number and timeliness of HMIS data submitted to the district medical officer's office	DHIS database
5	Health Financing	District health expenditure on recurrent costs and health development	PMORALG
6	Leadership and Governance	Number of supportive supervision visits to health facilities	District sources, SARA & Research data

NBS=National Bureau of Statistics; SPD=Sentinel Panel of Districts; SARA= Service Availability and Readiness Assessment; PMORALG=Tanzanian Prime Minister's Office for Regional Administration and Local Government; ANC=Antenatal Care; SP=Sulfadoxine Pyremethamine; TT=Tetanus Toxoid; DHIS=District Health Information System

Table 2.4: Indicators for FANC and EmOC programme results

No	FANC Indicators	Description/Definition	Source of Data	EmOC Indicators	Description/Definition	Source of Data
1	Antenatal care coverage	The percentage of women who utilised antenatal care provided by skilled birth attendants for reasons related to pregnancy at least once during pregnancy among all women who gave birth to a live child in a given time period	Data were collected from service delivery points and collated into the District Health Information System database	Availability of EmOC services	Ratio of EmOC facilities to population & geographical distribution of facilities	Data were extracted from the District Health Information database and from Service Availability and Readiness Assessment survey
2	ANC clients protected from Tetanus	Proportion of pregnant women receiving 2 or more Tetanus Toxoid injections during current pregnancy		Institutional delivery rate	Proportion of all births in population in health facilities	
3	ANC HIV testing rate	Proportion of pregnant women tested for HIV during current pregnancy		Caesarean section as a proportion of all births	Proportion of all births in population by Caesarean section in EmOC facilities	

Table 2.5: Contextual factors and their corresponding proxy indicators‡

No.	Main group	Type of context	Proxy indicators (with data for 2010 and 2011)
1	Demographic	Fertility Infant Mortality Under Five Mortality	Total Fertility Rate Infant Mortality Rate Under Five Mortality Rate
2	Socioeconomic	Female Literacy Levels	Proportion of 15-49 females with: 1. Adult and non-formal education 2. Primary education 3. Secondary education
3	Environmental	Water and Sanitation	Proportion of households using: 1. Drinking water from improved water sources 2. Improved toilet facilities
4	Infrastructural	Transport and Communication	Percentage of villages/streets in the district: 1. With accessibility to public (or own) transport services within 3 kilometres of district's road network 2. Whose roads were impassable during rainy season 3. Covered by the available telecommunication services and networks in the district
5	Natural disasters	Drought Floods Disease outbreaks Famine	Percentage of households in the district affected by natural disasters in 2010 and 2011: 1. Drought 2. Floods 3. Disease outbreaks 4. Famine
6	Public health system	Other non-FANC/EmOC programmes/interventions	Number of villages/streets or health facilities covered by interventions on: 1. Other maternal and newborn health programmes 2. Child health programmes (immunization, micronutrients, nutrition, etc.) 3. Malaria projects/programmes (indoor residual spraying, bed net use, etc.) 4. HIV/AIDS projects/programmes (Care and Treatment, Counselling and Testing, etc.) 5. Community health/Social protection programmes

‡Some of the data were obtained through asking district officials from within the District Executive Office (for example District Planning Officer, District Engineer, and others)

2.4 Design of research tools

Two questionnaires were developed for collecting primary data and one checklist was prepared for use as guidance in extracting data from secondary data sources. A maternal health experts' questionnaire was used to collect opinions from experts, district health authorities, and implementers of maternal health services on the weights they considered each programme component contributed to the strength FANC and EmOC programmes. This is appended as Appendix 1: "Maternal Health Experts

Questionnaire". The second questionnaire, "Data Collection Form for Contextual Factors in a District" was designed for data collection of the contextual factors in districts (Appendix 2). The checklist, "Indicator Guidance for Data Extraction from DHIS Database" was prepared to guide data extraction from a central database (Appendix 3). An additional questionnaire "Assessing Content and Face Validity of the Tool for Measuring Programme Implementation Strength" was also prepared for collecting expert opinions in determining the content and face validity of the approach used (Appendix 4). The service availability and readiness assessment (SARA) survey that generated some data for the programme implementation strength and for programmes coverage used its own questionnaires modified from those designed by the World Health Organisation. [137] These are not included in the appendices.

2.5 Training, logistics and data collection

There were two main groups of data collectors who collected most data for this study: 9 research assistants and 23 FBIS coordinators. In order to have good data quality, it was important to train all data collectors by giving them a comprehensive understanding of the tools, procedures, norms, and field logistics before data collection commenced. Each group of the data collectors were received training specific for their type of data they would collect. The nine research assistants collected data for the programme implementation strength and opinions from the maternal health experts. The 23 FBIS coordinators were separately trained as part of their full time job on collecting and managing the secondary data from health facilities in respective districts.

2.5.1 Research assistants

The nine research assistants were selected from a pool of data collectors from various projects run by the Ifakara Health Institute. The selection criteria were based on research assistants having a minimum qualification of a college degree and on having experience in data collection on similar fieldwork with the Ifakara Health Institute. The author prepared the training materials and trained the nine research assistants for two days. The training involved orienting the research assistants on the PhD overall goal, objectives and the research methods, a thorough review of the research tools, a reminder

on interviewing techniques, and adherence to local authorities' protocol. To have the research assistants gain the most from the training sessions, the author allowed in-session interactive questions and answers, general discussions and sharing of previous experiences. Both English and Swahili languages were used during training as data collection would involve a largely Swahili-speaking audience in rural and urban areas of Tanzania. As there was no formal assessment, it was hoped that by the end of training, research assistants had: acquired knowledge of and understood the questions in the questionnaires; gained the skills necessary to elicit good quality responses from respondents; had a common approach in conducting the survey; and that they understood the importance of adhering to ethical issues for interviewing human subjects.

The training was followed by piloting the questionnaires in Iringa Urban district (one of the 23 study districts). The team spent five days jointly collecting data. At the end of each day, the team convened at one location to discuss the success and challenges they encountered in collecting the required data from different groups of respondents. The team also reflected on how to manage time during interviews and discussed threats to data quality and their possible solutions. A few problems with questionnaires were identified and revisions were made based on piloting the instruments. From the pilot, research assistants were allocated two districts each. Each research assistant spent seven days in a district collecting data and moving between health facilities. After the two weeks were completed, six research assistants were further assigned to complete data collection in the remaining districts for seven more days.

2.5.2 FBIS coordinators

In 2010, Ifakara Health Institute recruited the 23 Facility-based Information System (FIS) Coordinators to be responsible for providing technical assistance in districts on health information management system and on data collection. Before they were deployed to their respective districts, FBIS coordinators attended a two-week training workshop on: how to collect and summarize data from health facility registers on a monthly basis, how to use the central database, the District Health Information System (DHIS) for data entry, how to capture and handle inconsistencies in data, how to

calculate completeness and accuracy of data from health facilities, and how to generate and transmit regular reports to the central office using the Internet. In addition, two data managers from the Ifakara Health Institute's central office also attended the two-week training to familiarize themselves with data pathways and quality issues from study districts. The two data managers were responsible for handling data from the coordinators through making regular queries in the central database and reporting and counterchecking with the coordinators for error correction and data quality.

2.5.3 SARA survey training

The Global Fund to fight Malaria, HIV/AIDS, and Tuberculosis provided funding to conduct the Service Availability and Readiness Assessment (SARA) survey through the Tanzanian Ministry of Health and Social Welfare that commissioned the work to Ifakara Health Institute. [SARA survey contributed a significant part of the data for the programme implementation strength and programmes coverage]. Training for the survey was conducted during the first quarter of 2012 and involved all the 23 FBIS coordinators and their 23 district counterparts – HMIS focal persons as well as senior staff from the Ministry of Health and Social Welfare and the Ifakara Health Institute. The author prepared the training and survey materials he acquired from the World Health Organization.[137] The senior staff provided the training to participants for three days at a central location and later visited all study districts to provide supportive supervision to the coordinators and to review data for completeness and quality.

2.6 Overview of methods

Table 2.6 shows the general overview of the methods with a linkage to the research objectives and data sources. To address each objective, methods used both primary and secondary data sources. The sections below provide a brief description of the methods. Detailed methods for each objective are later provided in respective chapters.

Table 2.6: Linkage between methods, research objectives and data sources

Research Objective	Methods	Data source/s
Objective 1: To develop an approach for estimating programme implementation strength by: developing a weighting scheme for estimating the strength of focused antenatal care and emergency obstetric care through use of opinions from maternal health experts; and applying the weighting scheme to programme indicators to estimate the implementation strength of the two tracer programmes in a nationally representative sample of 23 study districts.	The weighting scheme was developed by aggregating preference weights for each programme component (Chapter 3). Weights were obtained through self-administered questionnaires by maternal health experts. Programme data from study districts were obtained through surveys and access to official data sources (Chapter 4). Weights were applied to programme data to generate composite scores (overall implementation strength) – which were used to rank and compare the districts, by maps (Chapter 5).	Interviews of maternal health experts; Surveys and official district statistics.
Objective 2: To assess utilization and coverage of focused antenatal care and emergency obstetric care programmes in study districts.	Using proxy indicators, facility-based routine monitoring data (between Jan 2010 and Dec 2011) and some survey data were extracted two data sources (Chapter 6)	The District Health Information System database containing data from service delivery points, SARA survey.
Objective 3: To investigate and analyse district-level contextual factors as potential confounders of the effect of FANC and EmOC on programme utilization and coverage.	Summary measures of means, medians, standard deviations (and 95% confidence intervals) and differences between urban and rural districts were used to compare districts. (Chapter 7)	Investigation involved District Officials within the District Executive Officer to collect data for the contextual factors in respective districts
Objective 4: To test the programme implementation strength approach by investigating the association between programme implementation strength and utilization and coverage of the two programmes through an illustrative dose-response analysis	Univariable regression analyses and graphical methods were used to assess presence and strength of associations (Chapter 8)	Dose-response analyses involved data from Objectives 1, 2 & 3
Objective 5: To test the content and face validity of the programme implementation strength tools.	Percentages were used to represent overall content and face validity, and weighted Kappa statistics were used to assess the inter-rater agreement (Chapter 9).	Data from a follow-up survey of a sample of the experts.

2.6.1 Introducing the implementation strength method

The novel approach, "implementation strength for use in dose-response analysis" is designed to measure the strength of programme components and programme coverage changes over time by computing composite scores of constituent indicators as a reflection of their contributing effect in large-scale programmes such as FANC and EmOC. Programme components are those packages of inputs and activities that comprise a programme. Each programme component has its own set of proxy indicators that are part of the evaluation process. It is recommended that districts should be used in the evaluation process as units of design and analysis because of their central role in the health system. The evaluation method is packaged in a three-staged process, with each stage containing its own processes. Figure 2.2 shows the three stages of the implementation strength evaluation approach as being: evaluation planning, information gathering, and analysis – or the "PIA" stages in its short form ("P" for planning, "I" for information gathering and "A" for analysis).[f]

At the planning stage ("P"), the approach involves reviewing the literature and conducting an expert-opinion survey to identify the minimum essential indicators for: the implementation strength (resources/inputs and programme activities), the anticipated programme results, and for contextual factors. The concept of minimum essential indicators acknowledges that health system staff are over-worked and stressed and should therefore not be further burdened to collect even more data than they already do, but rather for them to focus on a few measures [essential set of indicators]

---

f The PIA acronym is not in reference to any other acronym in the literature and was not developed with any other approach in mind and should thus not be confused with other acronyms such as the "Privacy Impact Assessment" for collection, use or disclosure of personal information: http://www.isecurityconsulting.ca/pia-methodology-and-approach.php. Last accessed: 24th Dec 2014. It should rather be taken as a coincidence of acronyms. However, the three stages of planning, information-gathering, and analysis appear to be the essential ingredients of nearly all evaluations including programme cycle management of "plan-do-study-act cycles"(summarised here: http://pkpinc.com/files/NA01MoenNormanFullpaper.pdf), and more others. The novelty of the approach is therefore not in the PIA stages but in the implementation strength method for use in dose-response analysis of programme coverage changes over time.

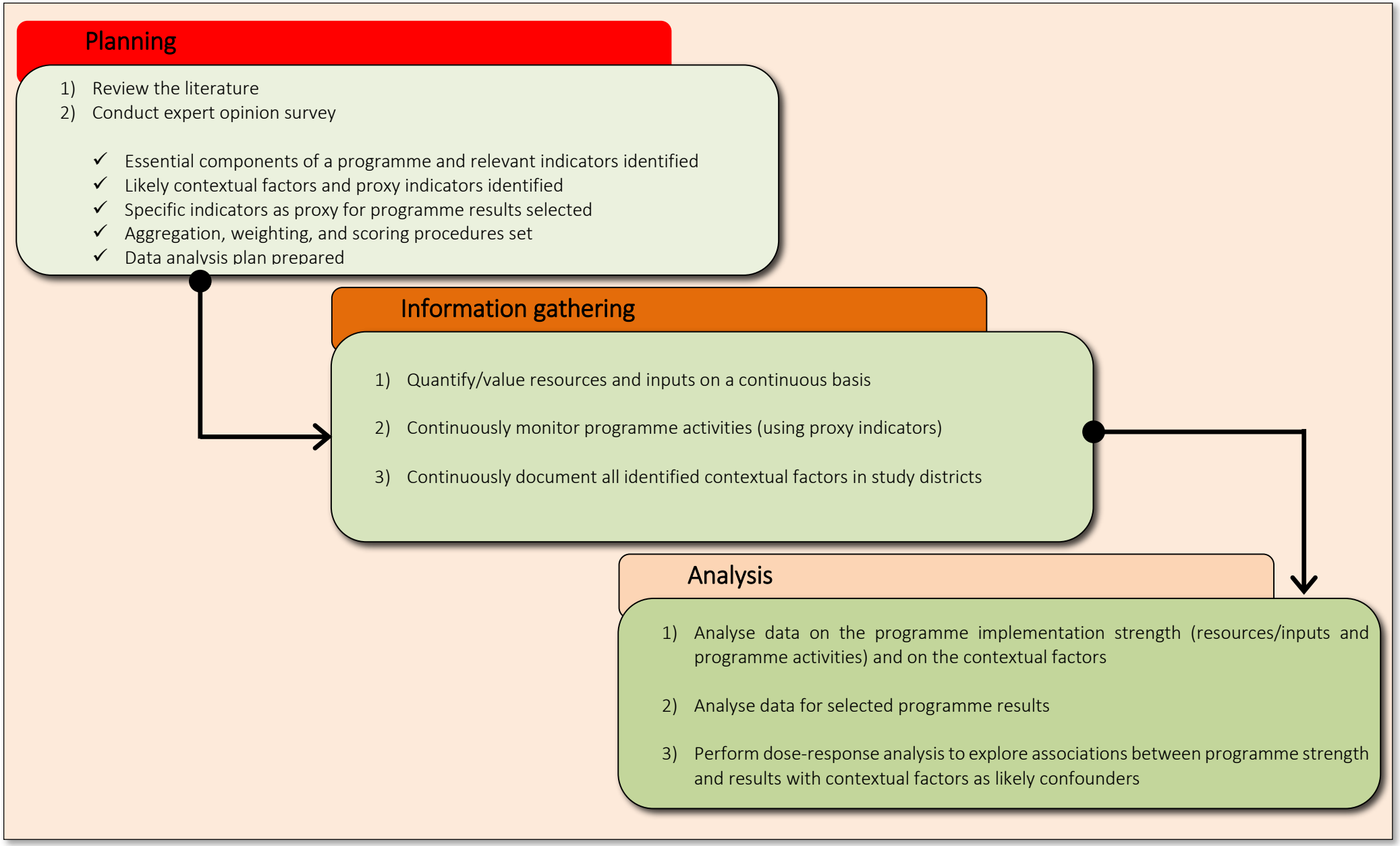


Figure 2.2: Stages of the implementation strength approach

that they can actually use themselves to help manage programmes as well as pass to national level, for example for use in implementation strength or other evaluation work. The planning stage should conclude by stating the procedures for aggregation, weighting, and scoring mechanisms as well as preparing the analytical plan before stage two is implemented.

During the information-gathering stage, programme evaluators will be required to continuously collect data using the identified minimal essential set of indicators by valuing all resources and inputs (where possible), monitor programme activities, monitor changes in contextual factors, and monitor programme results all along programme lifecycle. In stage three of the approach, at the end of a programme or at specific evaluation phase of a continuous programme such as the case of FANC and EmOC programmes), evaluators will be required to analyse all programme data collected from when the programme started to its end. By quantifying both programme activities/inputs and programme results, the approach provides room for evaluators to conduct a dose-response analysis through which, associations between programme efforts and programme results can be estimated. Dose-response analysis also offers the opportunity to adjust for the confounding contextual factors with which possible differences can be explained for districts having similar implementation strength but having varying programme results.

In applying the implementation strength approach, the conceptual framework (Figure 1.1) was prepared to guide the process and eventually test it with programme data. During the planning stage ("P"), the author conducted a review of the literature to identify a minimal set of possible indicators for use in evaluating the implementation strength, programme results, and the contextual factors. This was followed by the author conducting a survey of maternal health experts through which the weighting scales were obtained (more details are provided in Chapter 3). The author also prepared the procedures on how to aggregate and score the indicators and how to analyse each data category: for the implementation strength, for programme results and for the contextual factors.

For stage two of the approach, this study collected programme monitoring data for the two tracer programmes by conducting cross-sectional surveys and accessing data from other sources. Stage two needs three to four years of data to be able to conduct dose-response analysis of changes over time in stage three. Due to time limitations,

only single-point data were collected, mostly between January 2010 and December 2011. In addition to single-point data, the study did not include costing data due to timing and availability of funding. For the third stage of the approach ("A"), the author analysed all data categories including performing an illustrative dose-response analysis to show the approach's ability to assess associations between programme implementation strength and programme results. All three stages of the implementation strength approach are detailed in respective chapters.

2.6.2 Contribution of implementation strength in evaluations

The implementation strength concept can be used alone or to augment other designs. For example, in addition to tracking activities in areas with or without a given programme (that is, in the traditional before-and-after analysis with a comparison group) and in stepped wedge designs (in cases where a programme is implemented sequentially). Most importantly, the implementation strength approach has been designed in such a way that it provides programme evaluators with the opportunity to assess statistical associations between programme results and implementation strength adjusting for confounding factors. Results of the approach can potentially inform on how to improve programmes through dissemination of preliminary findings to government and partners. Besides its clear contribution in programme evaluations, the approach's requirement for continuous gathering of information (using proxy indicators) may be costly in settings where data collection is not integrated in established systems. For this reason, measurement of implementation strength could be of more relevance and of more direct use in districts with limited resources but where routine systems for data collections are well established.

2.6.3 The concept of minimal essential indicators

The concept of a minimal set of essential data might help to reduce the cost of data collection and improve data quality as responsible parties will have relatively fewer number of indicators to report on and more time to aggregate the data from different sources. Selection of indicators comprising the minimal essential set can be reached through reviewing the literature, identifying central government's priority

programme areas, and through in-country technical working groups and local programme managers and implementers.

The concept of minimal essential data has been used by several international aid organisations in evaluation of their programmes. For example, the World Health Organisation's Accountability for Women's and Children's Health had selected a set of only eleven indicators for monitoring maternal, newborn and child health, these were: [138] 1) Maternal mortality ratio 2) Under-five child mortality, with the proportion of newborn deaths 3) Children under five who are stunted 4) Proportion of demand for family planning satisfied (met need for contraception) 5)Antenatal care coverage (at least four times during pregnancy) 6) Antiretroviral (ARV) prophylaxis among HIV positive pregnant women to prevent HIV transmission and antiretroviral therapy for [pregnant] women who are treatment-eligible 7) Skilled attendant at birth 8) Postnatal care for mothers within two days of birth 9) Exclusive breastfeeding for six months (0–5 months) 10) Three doses of combined diphtheria-tetanus pertussis (DTP3) immunization coverage (12–23 months) and 11) Antibiotic treatment for suspected pneumonia.

Likewise, the Millennium Challenge Corporation and other donors had commissioned the Global Health Indicators Working Group (GHIWG) to prepare and recommend a set of indicators for use in evaluating a country's eligibility for investment in health.[139] GHIWG selected eight indicators for monitoring governments on health: 1) DTP3 immunization rate 2) Government public health spending 3) Under-five mortality rate 4) Stunting 5) Skilled birth attendants 6) Contraceptive prevalence rate 7) Unmet need for family planning and 8) Access to water. In general, the selection of minimal essential data should be reached in consideration of the indicators' value as proxy measures balanced against the likely data quality. Comprehensiveness and comparability of the indicators across different settings such as districts, regions or countries are also important.[140] Section 2.3.2 briefly discusses and presents the minimum set of indicators that were selected for use in the implementation strength approach for evaluation of the FANC and EmOC programmes.

2.6.4 Developing the weighting scheme

The study adopted the six World Health Organization's health-system-building blocks: human resources, drugs and supplies, service provision, health financing, health information systems, and leadership and governance, considering them as the essential components for the programme implementation strength of the two programmes of FANC and EmOC. The six blocks are synonymously referred to as programme components or programme implementation strength components. For each programme component, a list of constituent elements and activities was defined. A team of maternal health experts with experience in coordinating, implementing, or training on maternal health was selected. A total of 210 participants representing the national-level, the district-level and the health facility-level were asked to score each programme component considering their perceived contribution to the implementation strength of FANC and EmOC programmes. Scores (or preference weights) were averaged across all participants, results presented in tables, and graphs. The maternal health expert's opinions so derived were later applied to scale the weight of the actual programme data collected from study districts for both FANC and EmOC programmes. More details of the methods and the results of the opinions from the experts are presented in Chapter 3.

2.6.5 Unweighted programme data

A list of indicators was selected a priori for use in collecting programmes data from districts. As summarised in Table 2.3, different approaches were used to collect the data for the indicators of each programme component. The indicators for each programme component were:

1. Human resources – the number of health workers per 10,000 population (adopted from the World Health Organisation). Health workers included medical doctors, non-physician clinicians (including assistant medical doctors, clinical officers, lab technicians and pharmacists), and nursing and midwifery professionals.
2. Drugs and supplies – Availability of tracer drugs: antenatal care (SP drugs, tetanus toxoid injections, Iron, or Folic acid) and EmOC (oxytocin, ergometrine and magnesium sulphate).

3. Service delivery – Number of health facilities per 10,000 population, number distribution of inpatient beds per 10,000 population, and number of first antenatal care visits per 10,000 population.
4. Health information system – Quarterly HMIS reporting rate including the number of and the timeliness of HMIS data submitted to the district medical officer's office.
5. Health financing – per capita district revenues and expenditure.
6. Leadership and governance – Number of supportive supervision visits to health facilities.

For the human resources component, the numbers of health workers were obtained from the Ministry of Health and Social Welfare's official statistics for 2011. For the indicator's denominator, projected population levels were extracted from the Tanzanian National Bureau of Statistics. For the drugs and supplies component, data on the availability of tracer drugs and the essential supplies for FANC and EmOC services were obtained from the SARA survey. SARA survey also included data for the service delivery component's indicators. Health financing indicators' data were extracted from financial reports that were made available on the website of the Prime Minister's Office for Regional Administration and Local Government. To assess functioning of the health information system in study districts, data from health facilities were obtained from the central database with reporting and completeness of HMIS reports. For the leadership and governance component, a review of the district health management records was conducted by the FBIS coordinators to obtain the number of supportive supervision visits made by members of council health management teams to health facilities. A detailed account of the methods for each programme component and the score results for each study district are discussed in Chapter 4.

2.6.6 Weighted implementation strength scores

In order to determine the indicative strength of the tracer programmes in districts, this study applied the weighting scales on the implementation strength to generate weighted scores. The weighting scales were obtained through opinions from the maternal health experts on programme data collected from study districts. For each programme component within a district, the unweighted scores (described in section 2.6.2 above and presented later in Chapter 4) were first aggregated from indicators

data. The aggregated scores were then adjusted by applying the weighting scales to calibrate them against the views of the maternal health experts. Along the process, a sensitivity analysis was conducted to assess for the uncertainty in programme data from study districts, the likely biased opinions from the experts, and the overall reliability of the methods used to obtain the data (methods and results of the sensitivity analysis are detailed in Section 5.37 and 5.42 respectively). Using the resulting weighted scores, maps of Tanzania were produced to display the variations in the implementation strength among study districts using three colours representing the low-performing districts, the medium-performing districts, and the high-performing districts. More details on the methods and the results of weighted scores and maps are presented in Chapter 5.

2.6.7 FANC and EmOC utilization and coverage

Utilization and coverage of FANC and EmOC programmes in study districts was assessed using data that were extracted from two main data sources: the central database (District Health Information System) and the SARA survey. [Permission to access the data from the DHIS central database and SARA survey owned and maintained by the Ifakara Health Institute was given to the author by the Director of Research]. As summarised in Table 2.4, each programme had three proxy indicators that were used to assess their coverages in study districts. Indicators for the FANC programme were antenatal care coverage, HIV/AIDS testing rate among pregnant women, and proportion of pregnant women receiving two or more doses of tetanus toxoid injections. Indicators for the EmOC programme were availability of EmOC services, the rate of institutional deliveries, and the rate of Caesarean sections. Analysis of data from the six proxy indicators was conducted and tables of results were generated. Other than showing differences between rural and urban districts, no weighting was done for all coverage data. Chapter 5 gives details on data extraction and cleaning, and on the overall results regarding coverage of the two programmes in the study districts.

2.6.8 Contextual factors

Factors potentially affecting relevant health outcomes, including other maternal health programmes, were identified through a review of the literature (more details in Section 7.2.1). Resulting factors were further scrutinized in view of their relevance

to implementation, utilisation, coverage, and scale-up of FANC and EmOC programmes, and, the ease of obtaining factor-related data. As summarised in Table 2.5, the final list of the contextual factors included indicators on demographic, socioeconomic, environmental, infrastructural, natural disasters and other health programmes contextual factors related to FANC and EmOC. FBIS coordinators were provided with the questionnaire for the contextual factors and were asked to collect respective data by accessing official statistics from relevant departments in the districts or, where applicable and necessary, interviewing relevant key district staff. Data on the contextual factors were analysed and summary measures of means, medians, standard deviations (and 95% confidence intervals) generated. In recognising differences between urban and rural districts in implementing the tracer programmes, a further analysis was performed to compared districts with significance tests of the difference between rural and urban districts summarised using a two-sample Wilcoxon rank-sum. Indicators of contextual factors were later included in the illustrative dose-response analysis. Contextual-factor indicators were ultimately included in univariable regression analysis in the illustrative dose-response analysis as potential confounding variables. Chapter 7 gives the details of the methods and the results for the contextual factors.

2.6.9 Illustrative dose-response analysis

As highlighted in Section 1.8, in order to conduct dose-response analysis, it is essential that data for programme coverage (including data for the contextual factors) be collected before and after programme implementation. Availability of before-and-after data enables performing regression analysis that makes it possible to explore the relationship between implementation strength and coverage change, adjusting for confounders. Given that the coverage and contextual data were both limited to a single time point, a dose-response analysis was performed for illustration purposes only. The analysis involved univariable regression analyses and generating graphs. Univariable regressions were conducted to explore the nature of the relationships between individual programme components and the overall index of the programme results for each programme. Data for the predictor variables were drawn from the weighted composite scores of the six programme components of human resources, service delivery, drugs and supplies, health financing, health information system and leadership and governance, and the contextual factors of total fertility rate, infant mortality rate, under five mortality rate, female literacy

levels, water and sanitation, transport and communication, floods, drought, disease outbreaks and famine.

For the response variables, the overall index on FANC programme coverage was the mean score of antenatal care coverage, tetanus toxoid utilization, and HIV testing rate. The overall index on EmOC programme coverage was the mean score of the availability of EmOC services, institutional delivery coverage, and Caesarean sections. The respective predictor variables (programme components and contextual variables) showing significance at P=0.05 were selected for multivariable regression analyses. A table was prepared showing the final results of multivariable regression. This was followed by generating two graphs (one for each programme) for exploring the dose-response relationship between the overall programme implementation strength and the overall programme results. Chapter 8 gives the details of the illustrative dose-response analysis.

2.6.10 Content and face validity

In order to assess the suitability and acceptability of the questionnaires and approach for this study, the author selected a convenience sample of 50 raters from the 210 maternal health experts who were involved providing opinions for use in developing the weighting scales. The 60 raters included 10 national-level, 18 district-level, and 32 facility-level participants. All raters were asked to participate in the content and face validity study using a self-reported questionnaire, 12 months after their first contact with the research assistants who interviewed them for their opinions. To provide their responses, the national-level raters responded to the survey questionnaire through email communication to the author. FBIS coordinators were asked to distribute the survey questionnaires to the district-level and facility-level raters who were based in their respective districts. Coordinators collected back the duly filled-in questionnaires and entered the raters' responses in a web-based tool prepared by the author.

All questions in the rating instrument used a reversed 7-point Likert scale, with 1 = "strongly disagree" indicating that the rater was strongly against the suggested construct, and, 7 = "strongly agree" indicating that the rater was in complete agreement with suggested construct. Assessment of the overall content validity was based on the 'Maternal Health Expert' tool that was originally used to collect the

experts' opinions for developing the weighting scales. To evaluate its suitability, raters were asked on five parameters of overall algorithm, tool thoroughness, clarity of instructions, tool length, and tool scoring system. On the other hand, the overall face validity was assessed by asking the raters whether the 'Maternal Health Expert' tool was suitable for use in evaluating the perspective on programme implementation strength and whether raters would be willing to use the tool if they were in a position to do so. For each construct, percentages for each Likert point were calculated by dividing the number of participants choosing the point (1, 2, 3, 4, 5, 6 or 7) by the total number of participants. Weighted Kappa coefficients were used to assess the overall inter-rater agreement. More discussion and results on validity are presented in Chapter 9.

2.7 Data management

2.7.1 Primary data

Primary data here refers to programme data collected from study districts by the research assistants, which were not from secondary sources. To manage the primary data, an experienced data manager from the Tanzanian National Institute for Medical Research was hired to create databases, prepare data entry screens, and supervise data entry activities. All data entry screens were prepared using Epi-Data – a free package in Microsoft Windows environment, available from www.epidata.dk. For each data file created, the data manager created unique identification numbers of all entries to be able to manage duplicate records. Validation checks were written for a number of variables to allow valid entries in variables. For example, if 'facility_type' was used as a variable for entering the type of health facilities with likely values being 1=hospital, 2=health facility, 3=dispensary, and 9=other type, then validation checks could help prevent data entry errors for any invalid values outside those already specified.

Four data entrants double entered the data while working closely with and under the supervision of the data manager. In teams of two, the first data entrant entered the first entry while the second person entered the second entry of the same questionnaire. The data manager wrote and ran a program to merge the first and second entries. Whenever there was a mismatch between the two entries, the program appended the differences into a log file. Using the log file, the data manager

and data entrants corrected reported errors in discordant entries of the merged file with reference to the original forms.

2.7.2 Secondary data

Secondary data were all data that the study extracted from secondary sources. Secondary data were mainly those obtained from the SARA survey and from the DHIS central database containing facility-level data. Extracting data from the DHIS involved constant communication with two data managers responsible for data cleaning and management of health facility data, and with follow-up communications with district coordinators to verify particular data elements and entries.

2.8 Data completeness

Primary and secondary data collected from the 23 study districts had varying completeness depending on the type of tools used. Table 2.7 shows the names of study tools with their corresponding number of expected and actual data elements. Overall, 80% of respondents provided data for this study through the various study tools and questionnaires.

2.9 Ethical considerations

Data collection activities for this study received ethical clearances from both the London School of Hygiene and Tropical Medicine ethical committee and the Ifakara Health Institute Institutional Review Board. Ifakara Health Institute was already cleared by the Tanzanian National Institute for Medical Research for the FBIS arm of the Sentinel Panel of Districts – from which secondary data for this research were obtained. In addition, written informed consent forms were collected from all respondents who participated in providing primary data.

Table 2.7: Types and number of questionnaires before and after data collection, by objective

No.	Tool name / Questionnaire		For objective/s	Number of forms		% Completeness
				Expected	Actual	
1	Maternal Health Experts Questionnaire	National-level participants	**Objective 1:** To develop and test an approach for estimating programme implementation strength by: Developing a weighting scheme with opinions from maternal health experts for use on unweighted programme data	7	27	385%‡
		District-level participants		69	35	51%
		Facility-level participants		230	163	71%
2	Indicator Guidance for Data Extraction from DHIS Database†; SARA survey and other official statistics		Objective 1: Actual programme data collected to estimate the strength scores of the two programmes in a nationally representative sample of 23 study districts.	23	23	100%
3	Indicator Guidance for Data Extraction from DHIS Database †		**Objective 2:** To assess utilization and coverage of focused antenatal care and emergency obstetric care programmes in the 23 districts using 2-year routine monitoring data (Jan 2010 – Dec 2011).	23	23	100%
4	Data Collection Form for Contextual Factors in a District		**Objective 3:** To investigate and analyse district-level contextual factors as potential facilitators of the two tracer programmes FANC and EmOC to targeted populations.	Official statistics (23)	Official statistics (23)	100%
				Data from informants (23)	Data from informants (20)	87%
5	n/a		**Objective 4:** To test the approach by investigating the association between programme implementation strength, utilization, and coverage of the two programmes through an illustrative dose-response analysis.	n/a	n/a	n/a
6	Assessing Content and Face Validity of the method		**Objective 5:** To test the content and face validity of the method and tools used in order to explore the suitability of the approach.	60	50	83%
Overall Total				**458**	**364**	**80%**

FANC=Focused Antenatal Care; EmOC=Emergency Obstetric Care
‡More participants from nurse-midwifery colleges and maternal health programmes eventually agreed to participate beyond author's expectations
†Statistics for districts were extracted from the central database, DHIS=District Health Information System
SARA=Service Availability and Readiness survey. Objective 4 used data generated from Objectives 1, 2 and 4

Chapter 3: Developing the Weighting Scheme for Implementation Strength

3.1 Introduction

Due to the multifaceted nature of the health system in which large-scale interventions and programmes are implemented, obtaining evidence of the effect or impact of the programmes can be challenging. This is partly due to the challenges associated with obtaining reliable data from programme activities. Thus, effective tracking of implementation and efficient measurement of impact requires that programmes be rigorously designed and be continuously monitored and evaluated throughout the programme's life cycle. An essential part of tracking a programme is to identify, continuously document, and measure the key areas or programme components that are likely to cause the most impact.

Measuring programme implementation strength (defined as a quantifiable amount of efforts of a specific set of programme components and for a specific implementation time)[82 141] seems to be a plausible approach to evaluating large-scale programmes. By measuring and reporting the implementation strength, programmes are likely to recognize particular components that are mostly responsible in bringing about the programme effect. For example, a programme in Bangladesh implemented an intervention on the integrated management of childhood illnesses and conducted a midterm evaluation to assess its effectiveness on target populations.[142] Even though authors of the study did not use the term 'implementation strength', they found that three components were mainly responsible for the intervention's impact, and these were quality improvement in health facilities, increased use of facilities, and taking sick children to care providers.

Recently, the London School of Hygiene and Tropical Medicine, under the IDEAS project,[g] conducted a review of the literature on measuring implementation strength and concluded that there were no "rigorous strategies for measuring large-scale implementation of complex interventions in low income countries".[94] The review also found that most approaches were tailored to individual programmes and that there were three key methodological steps required in measuring implementation strength: 1) defining programme components, 2) outlining the procedures for developing measurement tools, and 3) providing a transparent approach for weighting and scoring systems. Acknowledging the methodological gaps in previous studies on measuring implementation strength of interventions and programmes, the current study adopted the approach of using the three key methodological steps for measuring implementation strength revealed in the review of the literature by the IDEAS project. This chapter describes how all three steps were achieved including the results of the data collected as each step was followed.

3.2 Methods

3.2.1 Design

This was an explorative study using a cross-sectional survey. The survey collected opinions from experts who were involved in maternal health services in various capacities and at different levels of the Tanzanian health system. Three groups of participants were asked to weigh (score) pre-selected programme components on their contributing effect to the programme implementation strength of FANC and EmOC programmes.

3.2.2 Sample selection

Between May 2012 and March 2013, the study recruited three groups of participants: national-level, district-level, and facility-level participants – each group having

---

[g] IDEAS (Informed Decisions for Actions) project is implemented by the London School of Hygiene and Tropical Medicine with funds from the Bill and Melinda Gates Foundation. Working in Ethiopia, North-Eastern Nigeria, and the state of Uttar Pradesh in India, the project's main aim is to improve the health and survival of mothers and babies through generating evidence to inform policy and practice. More information on the project can be accessed here: http://ideas.lshtm.ac.uk/

different levels of expertise and differing roles and experience. For the national-level group, participants had experience in either coordinating or managing maternal health programmes at the national level. In an effort to recruit national-level participants, the author obtained a list of members of a group called "Safe Motherhood Working Group" of the Tanzanian Ministry of Health and Social Welfare. The group's members had all the desirable criteria for inclusion into the study. Additional to the members of the Safe Motherhood Working Group, the study also recruited participants from universities and colleges involved in teaching and/or training nurses and midwives within the main city of Dar es Salaam. The study recruited a total 27 national-level participants.

The second group of participants included individuals with experience in planning and supervising maternal health programmes at the district level. In Mainland Tanzania, the "district"(or the "council" as it is famously referred to) is the primary administrative unit at which planning and implementation of policy-level decisions take place. Once every year, the Council Health Management Team meets to plan, budget, and prioritise all health-related activities including those for maternal and newborn health. In consideration of their roles as supervisors and coordinators of maternal health services in districts, and, as members of the Council Health Management Team, the District Medical Officer (DMO) and the District Reproductive and Child Health Coordinator (DRCHCo) and the Assistant DRCHCo, fulfilled the criteria for inclusion into the study as district-level participants. In total, the study recruited 35 participants from a pool of 69 potential participants (23 District Medical Officers, 23 DRCHCos and 23 Assistant DRCHCos).

Participants in the third group were drawn from health facilities. Health facilities are the first points of contact for clients in need of FANC and EmOC services. To be included in this group, participants were supposed to be staff members in charge of health facilities that provided reproductive and child health services. Persons in charge (or their assistants) are involved with supervision of day-to-day FANC and EmOC programme-related activities at the facility level. The study recruited 163 participants (at least 5 from each district) out of 230 originally intended participants from the 1,083 eligible facilities providing reproductive and child health services. For all three groups, the study recruited a total of 225 participants (national-level=27, district-level=35, and facility-level=163).

3.3 Experts' tool: design, piloting and administration

In order to identify and define the components of FANC and EmOC programmes (being the first of the three steps for measuring programme implementation strength), the author conducted a review of the literature by searching the PubMed database and Google search engines. Several related phrases were used such as 'program/me intensity', 'program/me strength', 'implementation strength' in an effort to find 'programme strategies' related to FANC and EmOC programmes.[75] [132] Following close assessment and comparison of the search results, the six WHO health-system-building blocks were selected to be the programme components for use in this study as core packages used for delivering FANC and EmOC programmes. The six building blocks are health workforce, drugs and supplies, service delivery, health financing, health information system and leadership and governance.[133] [134]human resources From the perspective of systems thinking, the six blocks do not work in isolation, but are connected by people who 'mediate', 'benefit' and 'act' in bringing about the effects within the health system.[134] Programmatically, the six building blocks constitute the essential functional parts that can be used to implement a large-scale health care programme to its target populations. This thesis refers to the six blocks as FANC and EmOC programme implementation strength components, and uses the term "programme components" interchangeably.

A questionnaire comprising the six programme components was prepared (Appendix 1: Maternal Health Experts Tool). For each programme component, a review of the literature was performed in order to select a list of the essential ingredients (elements/programme activities constituting a component). The Tanzanian Learner's Guide for ANC Service Providers and Supervisors and the WHO/UNFPA/UNICEF/AMDD Guidelines "Monitoring emergency obstetric care: a handbook" were the main documents used in reviewing the literature.[128] Other sources of information included (but were not limited to) the World Health Organisation website,[143] search of the PubMed database (with "maternal health" MESH term and the term "program"/"programme" in titles), and online resource materials from Jhpiego[144] and other global initiatives on maternal health programme. Preparation of Tanzanian FANC guidelines involved deliberations, consultative meetings and joint agreements between the Tanzanian Ministry of Health and the maternal health programme partners. The national guidelines were

officially launched in early 2000s for training providers and supervisors throughout Mainland Tanzania aiming at enhancing the provision of standard ANC services in an integrated manner. The guidelines were further reviewed and a new edition was released in 2008. Likewise, preparation of the EmOC handbook involved technical consultations from numerous interested parties spearheaded by global organisations of WHO, UNFPA, UNICEF, and AMDD.[128]

As an example, the search for the elements constituting the human resources component resulted in three elements: 'availability', 'competence', and 'productivity and responsiveness' (of doctors, nurses and midwives). Activities for the availability element were identified to include: enrolment of pre-service nurse-midwife students into allied health colleges, establishing and deploying Community Health Workers for maternal health care and life saving skills, allocating sufficient number of Anaesthetists, laboratory technicians, nurse-midwives and clinical officers in comprehensive EmOC facilities, and, distribution of sufficient numbers to all levels of health facilities. For the competence element, activities included: in-service training of service providers on FANC, EmOC and (IMPACC), practical training exposure to clinical officers and Assistant Medical Officers (AMOs) on EmOC in hospitals with specialists, continuing medical education and on-the-job training for all cadres, and, empowering traditional birth attendants (TBAs) through capacity building in disadvantaged/remote areas. The responsiveness and productivity element included: workplace motivation/incentives for doctors, nurses and midwives such as housing, hardship allowance, timely salaries and staff loans, payment for performance, and so on.

Figure 3.1 shows a modified Delphi process used to identify components' elements/activities and the steps followed in conducting the expert opinion survey. The author piloted the questionnaire by asking three national-level group members to review the proposed list of elements/activities for each programme component. The experts added 9 elements/activities to the tool and modified 3 of those originally suggested. The author revised the tool with the revisions from the experts and provided additional spaces to allow the remaining study participants (222) to also provide further modify and/or give comments on elements/activities of programme components. Table 3.1 shows the final list of activities/elements identified for all components of each programme.

The second key step in measuring implementation strength is to outline the procedures for developing measurement tools. The second step was achieved by preparing the "Maternal Health Expert Questionnaire" as the measurement tool. The questionnaire was designed for self-administration. An information sheet was attached as part of the questionnaire to help brief the experts on the study background and the reason for their involvement in the study. Nine research assistants were trained and later supported to distribute the questionnaires to both district-level and facility-level participants across all 23 study districts. The author distributed the questionnaires to all the 27 national-level participants. All completed questionnaires were collected from participants along with participants' written consent forms. Data were entered in a simple database and analysis performed in Stata Version 13 (Stata Corp, College Station, Texas, USA).

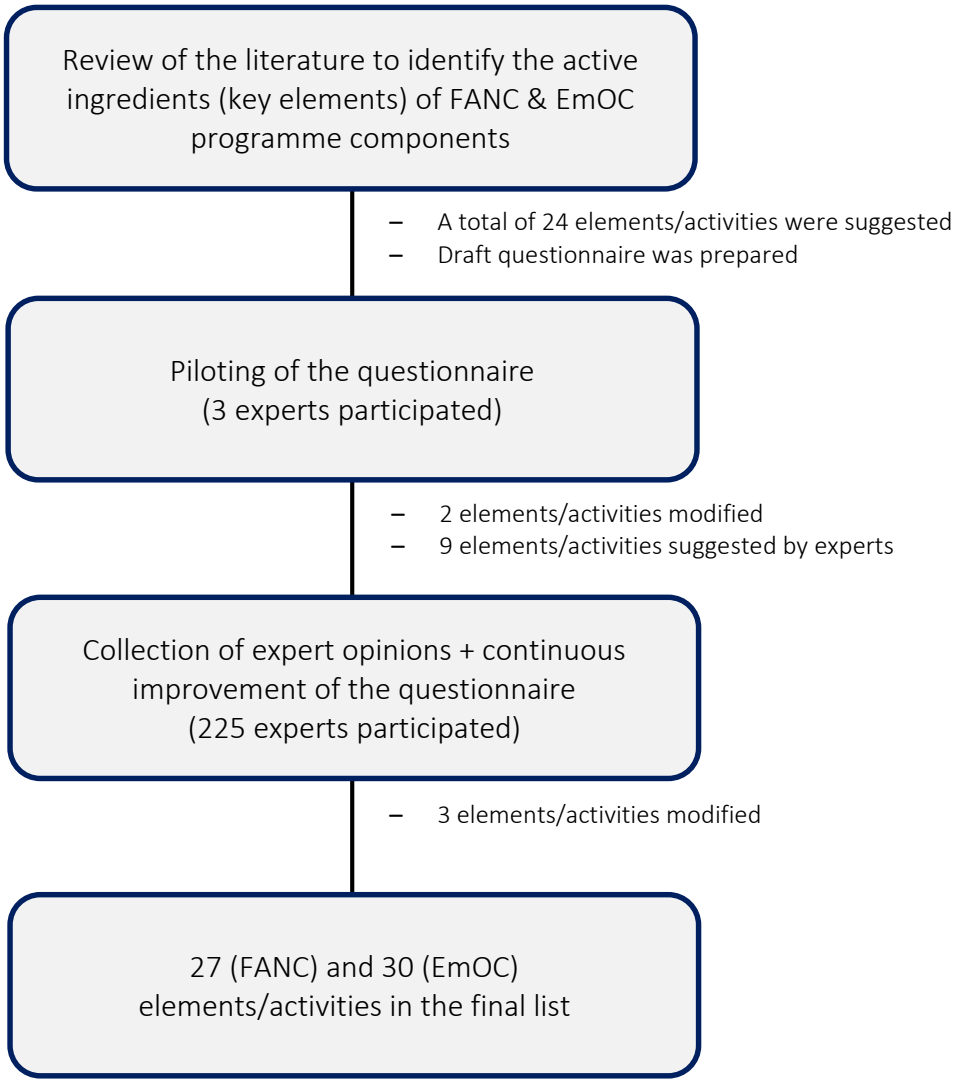


Figure 3.1: Modified Delphi process used to identify key elements/activities for each FANC and EmOC programme components

Table 3.1: Final list of programme activities/elements for programme components

Programme Component	FANC Programme Elements	EmOC Programme Elements
Human resources	1. Availability: a) Enrolment of pre-service nurse-midwife students in allied health colleges b) Establishing and deploying Community Health Workers for Maternal care c) Distribution of sufficient numbers to all levels of health facilities	1. Availability: a) Enrolment of pre-service nurse-midwife students in allied health colleges b) Allocation of sufficient number of Anaesthetists, Laboratory Technicians, Nurse-Midwives and Clinical Officers in Comprehensive EmOC facilities c) Distribution of sufficient numbers to all levels of health facilities d) Establishing Community Health Workers for Maternal care and Life Saving Skills
	2. Competence: a) In-service training of service providers on FANC b) Empowering TBAs through capacity building in disadvantaged areas	2. Competence: a) In-service training of service providers on EmOC/IMPACC b) Practical training exposure to Clinical Officers and AMOs on EmOC in hospitals with Specialists; Continuing Medical Edu and OJT for all cadres c) Empowering Traditional Birth Attendants through capacity building in remote areas
	3. Responsiveness & Productivity: a) Workplace motivation/incentives for HRH e.g. staff housing, hardship allowance, timely salaries and staff loans, monetary/non-monetary motivation (e.g. payment for performance), and so on	3. Responsiveness & Productivity: a) Workplace incentives for HRH e.g. staff housing, hardship allowance, timely salaries and staff loans, monetary/non-monetary motivation (e.g. payment for performance) and so on
Drugs and Supplies	1. Procurement of Essential (Tracer) Drugs for Antenatal Care	1. Procurement of Essential EmOC Drugs
	2. Procurement of Essential Equipment & Supplies for ANC	2. Procurement (and maintenance) of Essential Equipment & Supplies for EmOC
	3. Improving Integrated Logistics System (ILS) for drug supplies	3. Availability of emergency transport for patients; and for workers during emergency
		4. Improved Integrated Logistics System (ILS) for drug supplies
Service Provision (Incl. Infrastructure)	1. Improving ANC service provision (including PMTCT & FP)	1. Upgrading the existing health centres to enable them to provide both Basic and Comprehensive EmOC signal functions (building Operating Theatres, L&D wards)
	2. Referrals – within and to higher health facilities	2. Establishment of maternity waiting homes in hospitals
	3. Building new RCH Clinics; Renovating/regular maintenance of old/existing buildings	3. Allocating providers with houses w/in facility grounds for quick emergency calls
	4. Providing Incentive Vouchers (e.g. a pair of Kanga) to promote delivery in health facilities	4. Radio call communications & ambulances in selected health facilities
	5. Behavioural change communication programmes	5. Ongoing BCC programmes & community mobilization/empowerment initiatives
	6. Community male involvement	6. Outreach/Mobile services
	7. Community empowerment (e.g. accountability of service providers/CHMTs to community complaints, etc.)	7. Availability of Safe Blood
		8. Staff meetings
		9. Waste safe disposal
		10. Water and Power reliability (source, storage & emergency)
Health Financing	1. Increasing percentage of the Gov health budget spent on: • Salaries of health workers • Medicines and supplies • Other recurrent costs (e.g. admin cost at MOHSW, CHMT, In-service training)	1. Increased government health budget for expenditure on: • Salaries of health workers • Medicines and supplies • Other recurrent costs (admin cost at MOHSW, CHMT, In-service training)
	2. Increasing Health Insurance (Coverage, Funding, Policy) • Social Health Insurance – NHIF & CHF; Private for profit	2. Increasing Health Insurance (Coverage, Funding, Policy) • Social Health Insurance – NHIF; CHF; Private for profit
Health Information System	1. Availability: • Timely reporting of FANC information through quarterly reports from health facilities to the district	1. Availability: • Timely reporting of EmOC information through quarterly reports • Conducting maternal death audits
	2. Use: Incorporating routine information and evidence from research into district (and/or facility) for planning purposes on FANC	2. Use: Incorporating routine information and evidence from research (and maternal death audits) into district (and/or facility) for planning purposes on EmOC
	3. Capacity/Capability: • Purchasing and installing ICT equipment for data entry and data management in facilities/districts • Adequate number of registers & reporting forms in facilities	3. Capacity/Capability: • Purchasing and installing ICT equipment for data entry and data management in facilities/districts • Adequate number of registers and reporting forms in facilities
Leadership/ Governance	1. National Level advocacy for maternal health	1. National Level advocacy for maternal health
	2. Supportive supervision visits to health facilities	2. Supportive supervision visits to health facilities
	3. Issuance of 1s, rules and regulations – licensure of health professionals, accreditation & certification	3. Issuance of guidelines, Charts and Algorithms (Wall Charts), rules & regulations
	4. Development by Decentralization (D4D)/Basket Fund use Vs Local Gov funding policy	4. Development by Decentralization (D4D)/Basket Fund use Vs Local Gov funding
	5. Involving health and non-health stakeholders in defining and prioritizing health needs and services at the national level	5. Involvement of health and non-health stakeholders in defining and prioritizing health needs and services

The third key step in measuring the implementation strength of a programme is providing a transparent approach for weighting and scoring. To achieve the scoring system, a question was added in the Maternal Health Expert Questionnaire asking participants to score each of the six programme components on a scale between 0% and 100% according to what they considered each component contributed to the strength of a programme (FANC or EmOC). All scores were supposed to sum to 100%. Further, a second question was included asking participants to rank elements/activities within each programme component of each programme in the order of their importance/priority. For example, if a programme component had five elements/activities, experts were therefore supposed to rank the elements from first to fifth. Both the use of percentages for scoring programme components and priority-ranking of the elements/activities were tested during the pilot phase to determine whether it was easy for participants to perform both schemes. The pilot revealed that the scoring and ranking of elements/activities was clear to all three participants. However, one of the participants in piloting the tool over scored the human resources component by assigning it with 80% and the remaining 20% across the other five components. Curious on this, the author followed up with the participant and found that this person was once a senior government official involved with a nationwide project on human resources. Use of scores from 0% - 100% has been used in previous studies, with two particular studies using the method with maternal health experts in 49 countries[98] and 55 countries.[107]

The scores provided by the maternal health experts were averaged to provide the weighting system part of the third step. For each programme (FANC or EmOC), the weight of a component was equal to the mean of all scores provided by participants for that component. The mean was calculated by adding all scores within the component dividing the sum by the number of experts who scored the component. For example, if the health financing component for the FANC programme was rated by 206 participants whose sum of their scores was 3,770, then, the health financing component for FANC programme would have a mean score of 18.3% (3,770 ÷ 206) – as its overall contribution to the strength of FANC programme. Weights of programme components obtained through this process were later applied to the results from actual programme data (Chapter 5) to determine each component's 'adjusted' contribution to programme implementation strength.

3.4 Data analysis

3.4.1 Missing values

There were three records with missing values on FANC programme weights. Twelve other records contained weights that did not add to 100% as required. All 15 records (7%) with either missing or erratic values were excluded from analysis. For EmOC programme components, eleven records (5%) had missing values across all six components and were excluded from the analysis. The 15 records excluded from FANC programme analysis were not necessarily the same as the 11 records excluded from the EmOC programme analysis.

3.4.2 Experts' preferences

Opinions from the maternal health experts were expressed in percentage weights. These were analysed in Stata (version 13, Stata Corp, College Station, Texas, USA). Standard deviations and confidence intervals were calculated to assess variability of assigned weights. One-way analysis of variance was used to determine if there was any evidence of differences between the means of components provided by national-level, district-level and facility-level groups of experts. If a difference between groups was detected at $P<0.05$, multiple comparison tests were performed to disclose the pair causing the differences.

3.5 Results

3.5.1 Participants

The study intended to reach at least 7 national-level participants, 3 district-level participants from each district (a total of 69 from the 23 study districts) and at least 10 facility-level participants from each district (a total of 230 participants from the 23 district, with each participant from a different facility in the district). There was no pre-specified target number of participants (sample size) for the expert opinion survey. However, the study aimed at recruiting participants from all three levels of the health system in Tanzania including the national-level, the district-level and the health-facility-level. As shown in Table 2.7, a total of 27 national-level participants were enrolled to participate in the study over and above the 7 anticipated

participants. The surge from 7 to 27 was mainly due to participants from medical universities volunteering to participate in the expert opinion survey. For the district-level participants, the study recruited 35 of the intended 69 depending on the availability of the participants on the survey day. The study recruited 163 out of the 230 anticipated participants from health facilities – also depending on their availability on the survey. All national-level participants were urban-based while the majority of those from district-level and facility-level participants were based in rural settings.

Of the 225, 210 respondents had completed the survey questionnaire on the section for FANC programme as instructed. Table 3.2 shows that 13% (n=27) were national-level, 13% (n=28) were district-level and 74% (n=155) participants were from health facilities. Table 3.2 also shows that about 60% of participants were Clinical and Nursing Officers whereas Medical Officers, District RCH Coordinators, Directors, and Senior Officers comprised 13%, 9%, and 5% of the respondents respectively. Programme Managers and Officers, District Medical Officers and FANC/EmOC Teachers and Trainers were 4%, 3%, and 3% respectively. Participants with other designations constituted 4% of total respondents. There was no targeted sampling of participants based on their cadres but as the results show, most of the participants in health facilities were medical doctors, clinical officers and nursing officers whereas participants in other levels were either doctors or nurses even they preferred to be identified by their titles, such as district RCH coordinators and FANC/EmOC trainers and directors or senior officers.

Table 3.2: Experts participation by designation and health system level

Participant's Designation	Health System Level			Total Number (%)
	National	District	Facility	
Clinical Officers	0	0	64	64 (30%)
Nursing Officers	1	4	57	62 (30%)
Medical Officers	0	0	27	27 (13%)
District RCH Coordinators	0	18	0	18 (9%)
Directors/Senior Officers	10	0	0	10 (5%)
Programme Managers/Officers	9	0	0	9 (4%)
District Medical Officers	0	6	1	7 (3%)
FANC/EmOC Teachers or Trainers	7	0	0	7 (3%)
Other Designations	0	0	6	6 (3%)
Total number (n)	27	28	155	210

3.5.2 Components' contribution to FANC strength

Frequency distributions and normal density graphs shown in Figure 3.2 indicate that component weights assigned by the 210 expert participants for the FANC programme were normally distributed, with some of the components showing skewness with extended tails to the right. The distributions also indicate that there was some heaping of the weights at 10%, 20%, and 30%. It is possible that respondents had similar views of weights around components or heaping could also been due to some drawbacks of the scoring system – even though this was not verified. Box plots shown in Figure 3.3 also show how outlier weights might have deflected medians from means, resulting in skewness of assigned weights, especially those for the human resources and health information system components whose median values are shown to be equivalent to or just above the first quartiles.

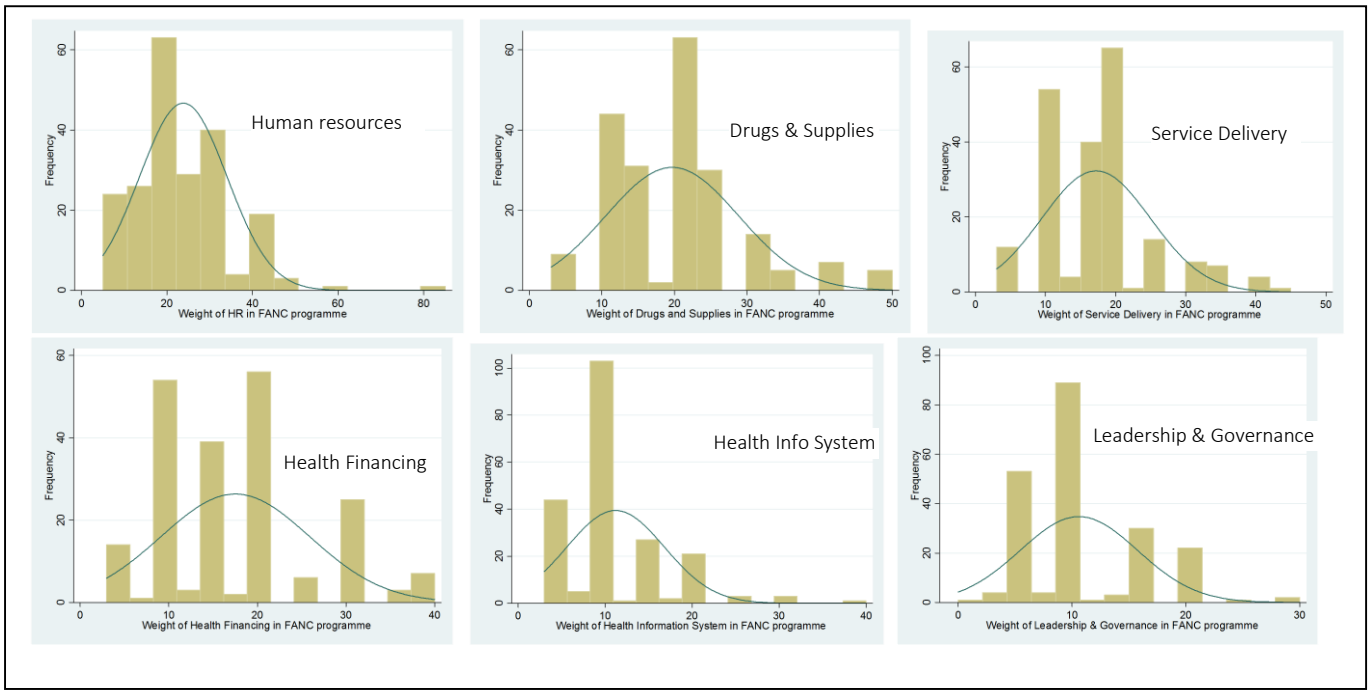


Figure 3.2: Frequency distributions (& normal densities) of FANC components (n=210)

Figure 3.4 shows the estimates (expressed in mean percentages) of what each component was thought to contribute to FANC implementation strength. The figure shows national-level, district-level and facility-level estimates with a summary of overall component means, their associated standard deviations, and, 95% confidence intervals. Generally, the 210 experts considered the human resources component to

have a greater weight (mean=23.8%) than other components. The 210 experts weighted the human resources component slightly more than double the weights of health information system (mean=11.2%) and leadership and governance (mean=10.6%). Drugs and supplies component came in second place with an overall weight of 19.7%, while health financing component contributed 17.5% and service provision contributed 17.2% to the strength of implementing the FANC programme.

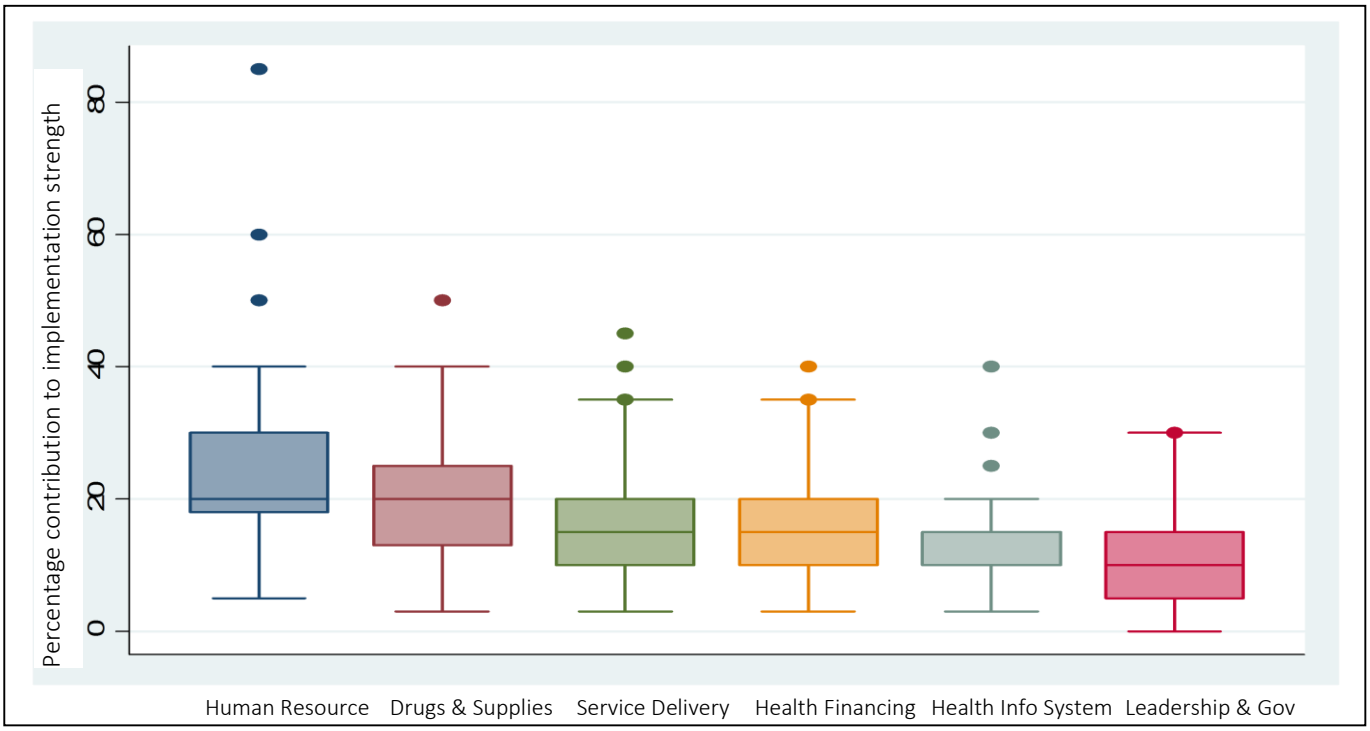
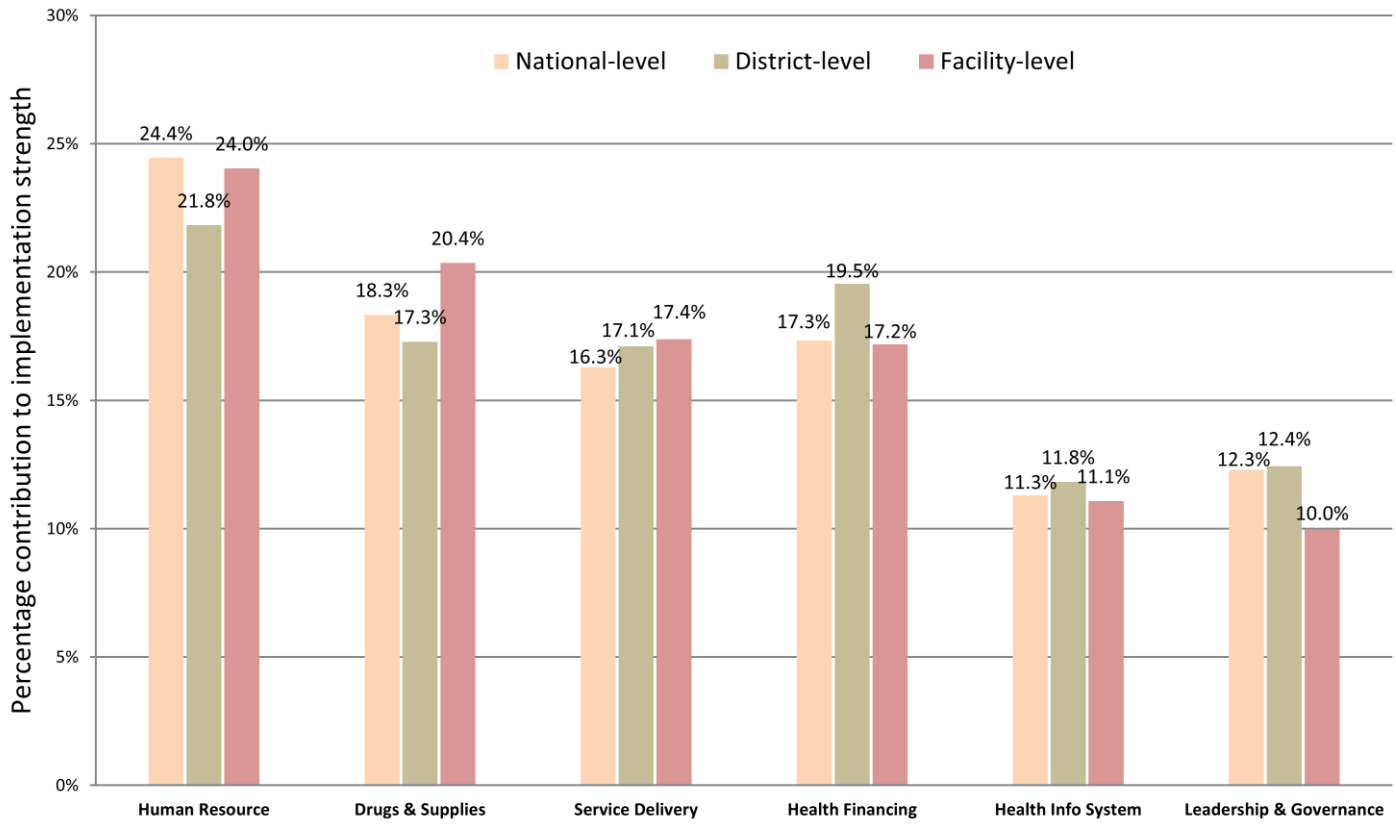


Figure 3.3: Box plots of percentage weights of FANC components (n=210)

Opinions from the experts in all three levels were similar for the component of service delivery for which national-level participants assigned 16.3% slightly lower than district-level (17.1%) and facility-level (17.3%). For the health information system component, national-level participants assigned a weight of 11.3%, similar to weights assigned by district-level participants (11.8%) and by facility-level participants (11.1%). On the other hand, facility-level experts assigned more weight to drugs and supplies (20.4%) compared to weights assigned by national-level (18.3%) and district-level experts (17.3%). Likewise, district-level participants assigned more weight to health financing (19.5%) as opposed to 17.3% and 17.2% weights assigned by national- and facility-level participants respectively. Notably, national-level and district-level participants assigned more weight to the leadership and governance component – 12.3% and 12.4% respectively, compared to 10.0%

assigned by facility-level participants. Even though the 95% confidence intervals were narrow (implying reasonable precision), there were notable variations in assigning weights to programme components. Variations in weights of the human resources component (standard deviation=10.2%) and of drugs and supplies component (SD=9.2%) were slightly higher than those of health financing (SD=8.4%) and service delivery (SD=7.8%).



	Human Resource	Drugs & Supplies	Service Delivery	Health Financing	Health Info System	Leadership & Governance
Observ (n)	210	210	210	210	210	210
Overall	23.8%	19.7%	17.2%	17.5%	11.2%	10.6%
SD	10.2%	9.2%	7.8%	8.4%	5.6%	5.2%
[95% CI]	22.4%	18.4%	16.1%	16.4%	10.4%	9.9%
	25.2%	20.9%	18.3%	18.7%	12.0%	11.3%

Figure 3.4: Mean percentage contribution of FANC components to implementation strength, by participant level, with overall means, standard deviations (SD), 95% CI

Variation in assigning weights of the health information system and leadership and governance components were relatively smaller, with SD=5.6% and SD=5.2% respectively. Bigger variations observed in the weights assigned to the human resources and drugs and supplies components are explained by the wide range of weights across the 210 participants, from a minimum of 5% to a maximum of 85% for human resources, and from a minimum of 3% to a maximum of 50% for drugs

and supplies. Tests for the differences between the level means are presented in Table 3.3. The observed differences between mean weights assigned by participants at the three levels were not significant at the 0.05 level, except for the leadership and governance means. The difference observed on the means for the leadership and governance component (P=0.012) – although after adjusting for multiple hypotheses, was not significant.

Table 3.3: Analysis of variance of means, participants' views, FANC programme

Component	F-test score	Overall p-value of differences	Bonferroni's p-values of the differences in level pairs		
			Facility Vs District	Facility Vs National	District Vs National
Human resources	0.620	0.541	0.882	1.000	1.000
Drugs & Supplies	1.680	0.189	0.310	0.870	1.000
Service Delivery	0.220	0.799	1.000	1.000	1.000
Health Financing	0.940	0.391	0.518	1.000	0.998
Health Info System	0.210	0.810	1.000	1.000	1.000
Leadership & Governance	4.490	0.012	0.059	0.089	1.000

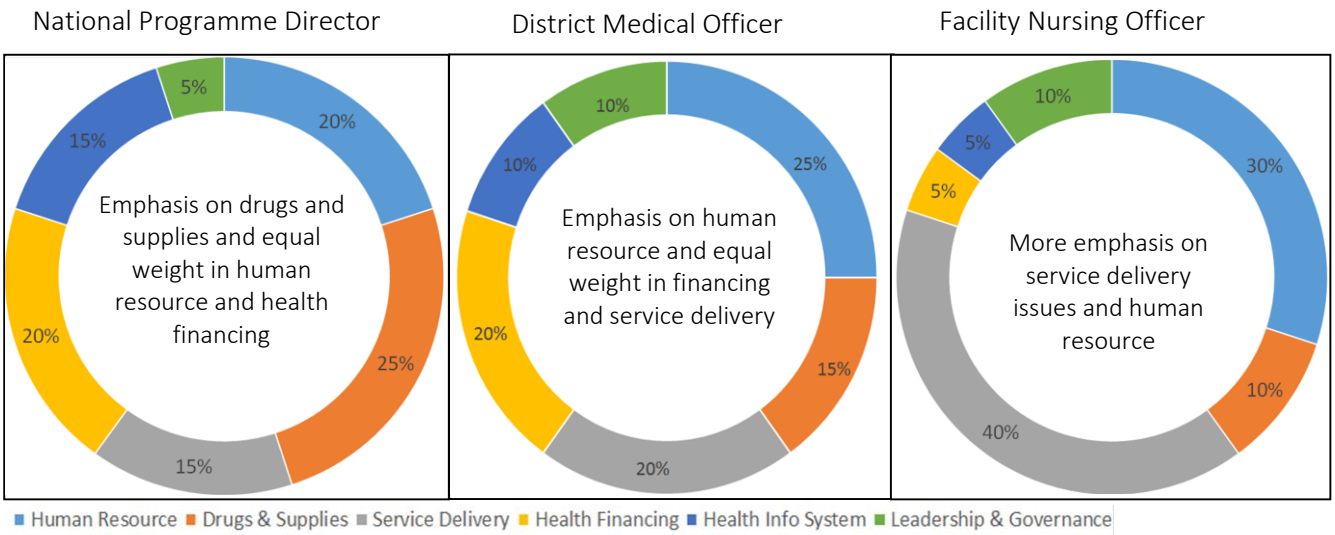


Figure 3.5: An example of how participants assigned weights to individual components from their experience- or level-based opinions

Figure 3.5 shows an example of how participants assigned weights for the individual components by their levels. The figure shows that national-level participants tended to put emphasis on drugs and supplies and equal weight in human resources and health financing whereas district-level participants tended to assign more weights on human resources with equal weights between health financing and service delivery. On the other hand, health facility participants seemed to assign more emphasis on service delivery issues and human resources.

3.5.3 EmOC programme implementation strength

Similar to FANC programme components, frequency distributions and normal density graphs for EmOC programme components shown in Figure 3.6 indicate that assigned weights were normally distributed with some skewness to the right. Box plots in Figure 3.7 show how outlier weights might have deflected medians from means, resulting in skewness of assigned weights especially those for the service delivery and leadership and governance components, whose median values are equivalent to or just above the first quartiles.

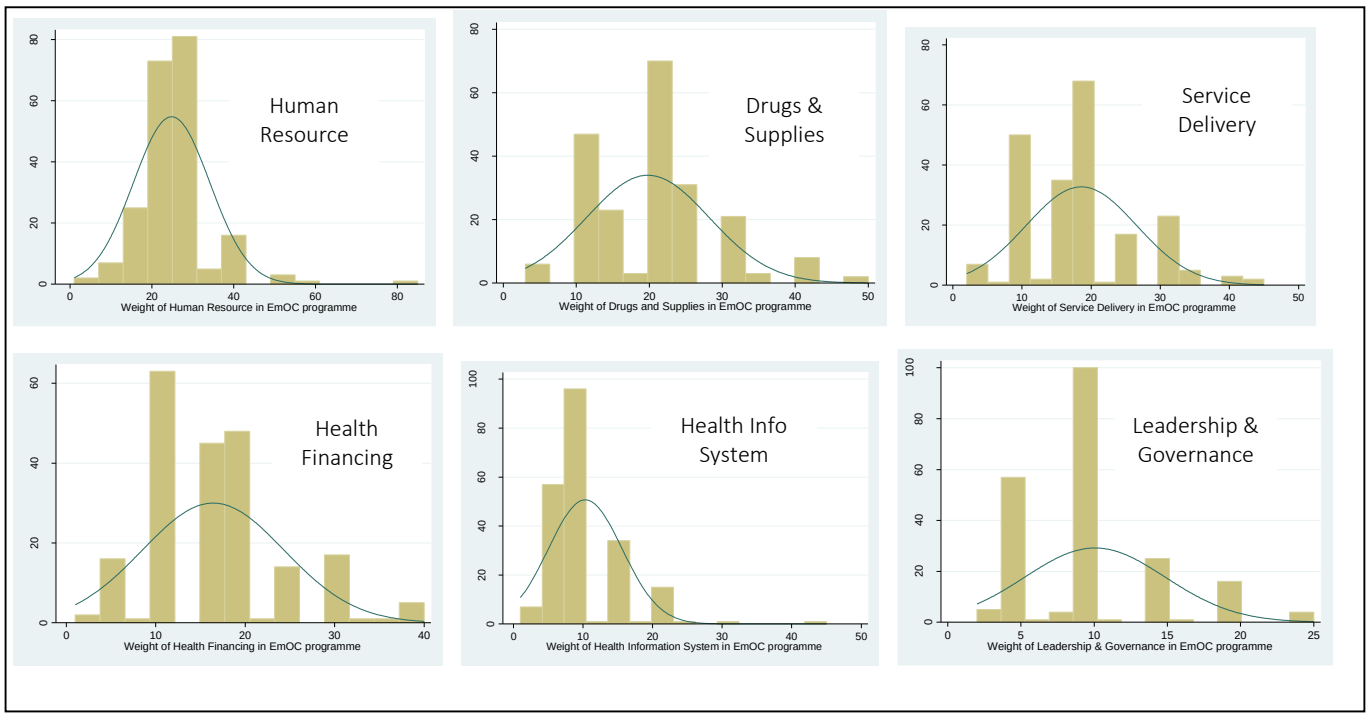


Figure 3.6: Frequency distributions (& normal densities) of EmOC component weights

Figure 3.8 shows national-level, district-level and facility-level estimates with a summary of overall component means and their associated standard deviations, and 95% confidence intervals. Similar to the FANC programme, experts assigned much more weight to the human resources component than to other programme components. The mean weight assigned to the human resources component (about 25%) was nearly two-and-half times the mean weights of health information system and leadership and governance components that contributed about 10% each. Experts assigned a weight of about 20% to drugs and supplies component while assigning about 19% and 16% to service delivery and health financing components respectively.

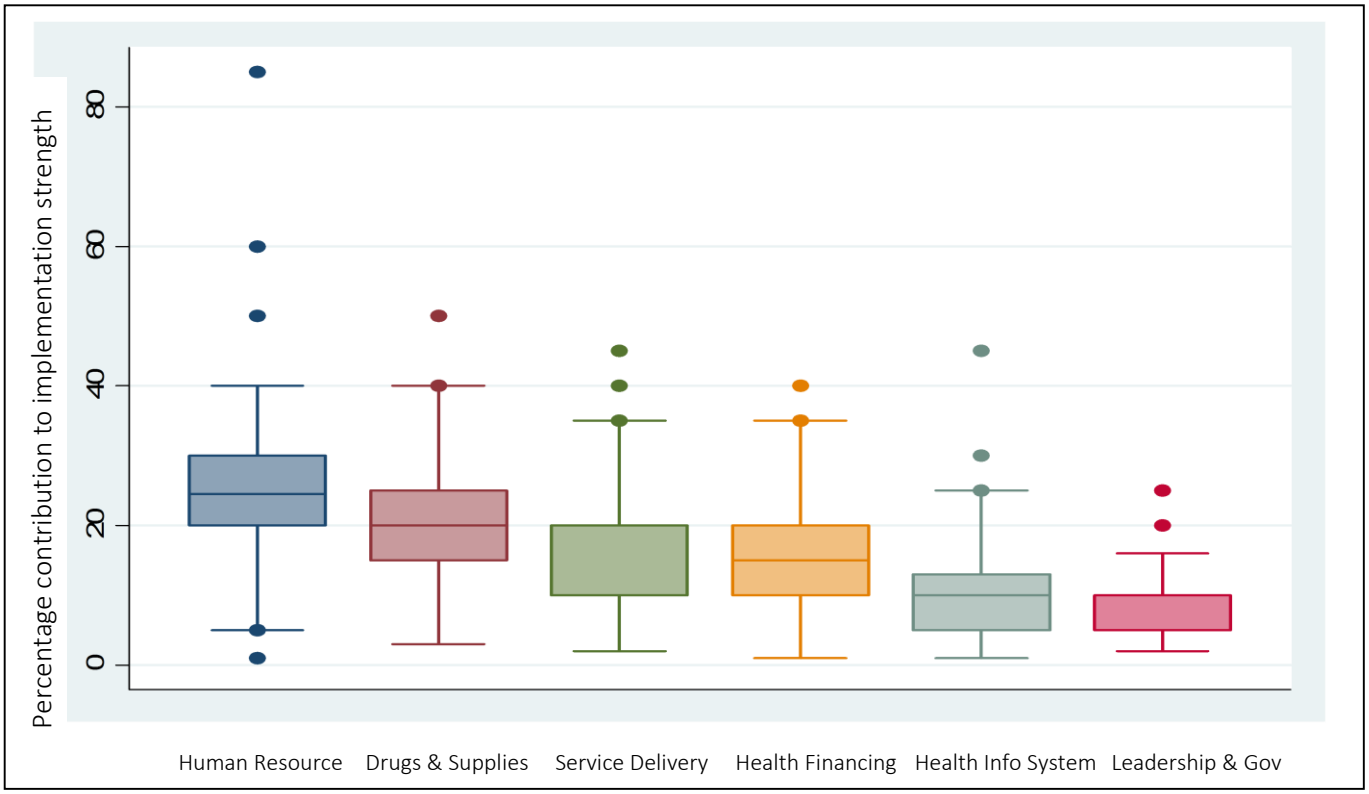
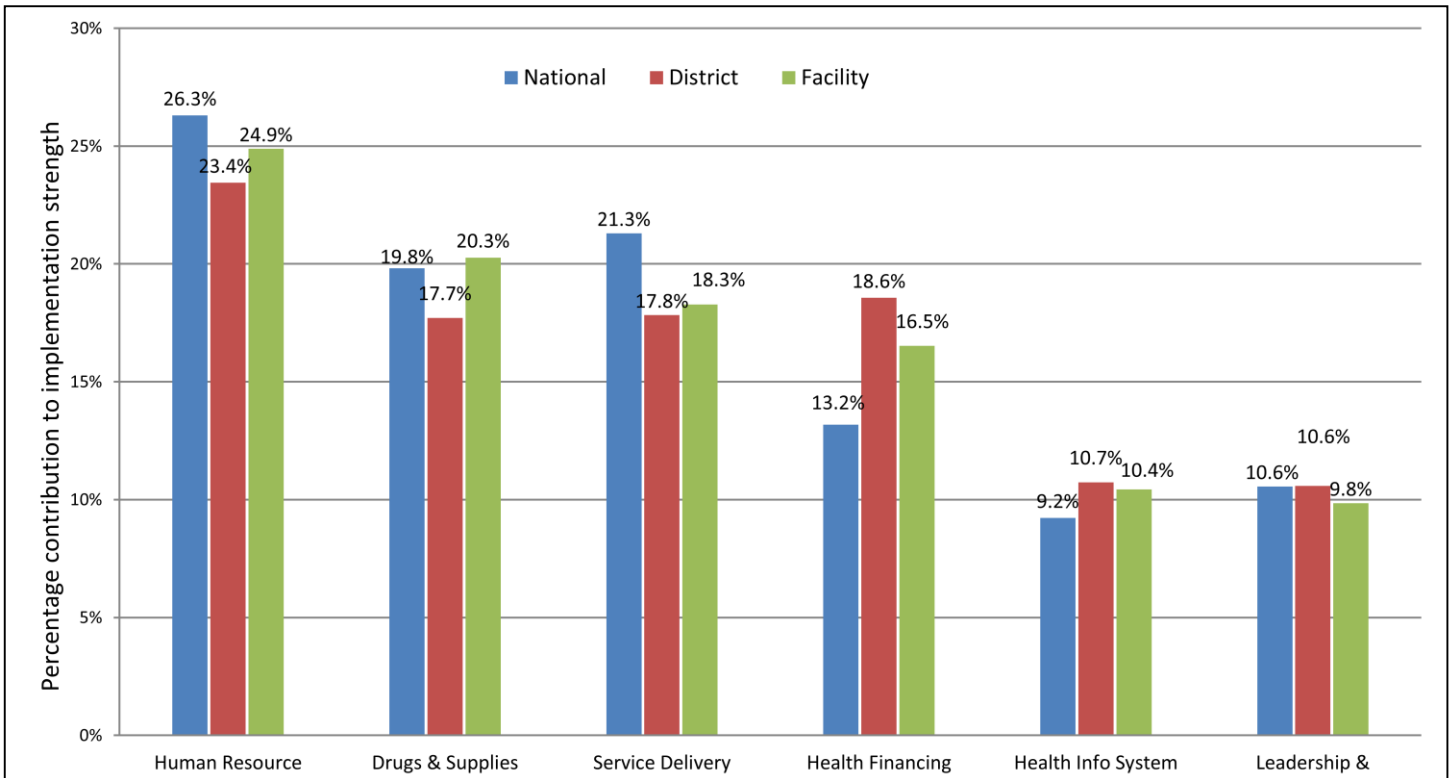


Figure 3.7: Box plots of percentage weights of EmOC components (n=214)



	Human Resource	Drugs & Supplies	Service Delivery	Health Financing	Health Info System	Leadership & Governance
Observ (n)	214	214	214	214	214	214
Overall Mean	24.8%	19.8%	18.6%	16.4%	10.3%	10.1%
SD	9.4%	8.4%	8.0%	7.9%	5.3%	4.8%
[95% CI]	23.6%	18.7%	17.5%	15.4%	9.6%	9.4%
	26.1%	20.9%	19.7%	17.5%	11.0%	10.7%

Figure 3.8: Mean percentage contribution of EmOC components to implementation strength, by participant level, with overall means, standard deviations (SD), 95% CI

National-level experts assigned more weight to the human resources component compared to their district-level and facility-level counterparts: 26.3%, 23.4%, and 24.9% respectively. Likewise, national-level experts assigned 21.3% to the service delivery component, much higher than the weights of 17.8% and 18.3% assigned by district- and facility-level participants respectively. Similar to the FANC programme, district-level experts assigned more weight to the EmOC health financing component compared to national- and facility-level participants: 18.6%, 13.2%, and 16.5% respectively.

On the other hand, experts in all three levels assigned nearly the same weights for the leadership and governance component, with weights of 10.6%, 10.6%, and 9.8% by national-level, district-level, and facility-level respectively. National-level experts assigned a little less weight to health information system: 9.2% compared to district- and facility-level participants who weighted the component 10.7% and 10.4% respectively. There were some variations between experts in assigning weights to programme components, although the 95% confidence intervals were narrow. Variation among experts was slightly higher in assigning weights to the human resources component (SD=9.4%) compared to variations for the drugs and supplies component (SD=8.4%), service delivery component (SD=8.0%) or health financing component (SD=7.9%). Variations in assigning weights were relatively smaller for the health information system and leadership and governance components, with standard deviations of 5.3% and 4.8% respectively. Table 3.4 shows that the differences between levels only for the health financing component means (P=0.03), which, were due to the difference between district-level and national-level participants (Bonferroni's P=0.025).

Table 3.4: Analysis of variance of the means, participants' views, EmOC programme

Component	F-test score	Overall p-value of differences	Bonferroni's p-values of the differences in level pairs		
			Facility Vs District	Facility Vs National	District Vs National
Human resources	0.710	0.494	1	1	0.715
Drugs & Supplies	1.290	0.278	0.330	1.000	0.996
Service Delivery	1.830	0.163	1.000	0.213	0.278
Health Financing	3.580	0.030	0.596	0.128	0.025
Health Info System	0.720	0.489	1	0.827	0.808
Leadership & Governance	0.490	0.611	1	1	1

3.5.4 Priority ranking of FANC activities

Table 3.5 indicates that, on average, experts on all three levels ranked the availability of human resources as top priority for FANC programme activities over and above competence or responsiveness and productivity. For activities related to drugs and supplies, district- and facility-level experts prioritized procurement of tracer drugs for antenatal care and improving the logistic system for supplying the drugs higher than the importance of procuring essential equipment and supplies for antenatal services. On FANC programme activities related to provision of service, experts prioritized improvement in provision of antenatal services including prevention of mother to child transmission of HIV (PMTCT), family planning and enhancing referral services within facilities and to higher facilities. Construction of new reproductive and child health clinics and renovating existing buildings followed in the priority of activities.

For the health financing component, experts gave more priority to increased government health budget for salaries and medicines than to increased coverage, funding, and policy for health insurance. Experts also agreed on priority of availability, use, and increased capacity of health information over district's increased capacity for information, communication, and technology (ICT) equipment including forms and registers for reporting health data. Supportive supervision visits to health facilities by district health personnel was given much more priority by the national- and district-level participants compared to their facility-level counterparts. Involvement of health and non-health stakeholders in setting priorities during planning was, on average, the second priority, with issuance of guidelines, charts/algorithms, rules and regulations and government decentralization by development being the lowest priorities.

Table 3.5: Facility-level means priority rank of activities/elements within FANC components, by levels

FANC Programme Component	Activity/Element	Priority rank of activities within components			
		National	District	Facility	Overall
Human resources	Availability: increased nurse-midwives pre-service students, etc.	1.8	1.7	1.9	1.8 (1st)
	Competence: in-service FANC training for providers	1.9	2.2	1.9	2.0 (2nd)
	Responsiveness & Productivity – provide incentives to staff	2.3	2.1	2.2	2.2 (3rd)
Drugs & Supplies	Procurement of Essential (Tracer) Drugs for Antenatal Care	2.2	1.8	1.9	1.9 (1st)
	Procurement of Essential Equipment & Supplies for ANC	1.7	2.1	2.3	2.2 (3rd)
	Improving Integrated Logistics System (ILS) for drug supplies	2.1	2.1	2.1	2.1 (2nd)
Service Provision	Improving ANC service provision (including PMTCT & FP)	2.1	2.1	1.9	2.0 (1st)
	Referrals – within and to higher health facilities	3.3	2.8	3.6	3.5 (2nd)
	Building new RCH clinics and renovating existing buildings	4.3	4.4	3.9	4.0 (=3rd)
	Providing incentive vouchers to attract women to deliver in facilities	6.3	5.5	5.0	5.2 (6th)
	Conducting behavioural change communication programmes	3.4	4.6	4.3	4.3 (4th)
	Increase community male involvement	4.5	3.5	4.1	4.0 (=3rd)
	Empowering communities-e.g. accountability of providers to community	3.6	5.0	5.4	5.1 (5th)
Health Financing	Increased government health budget for salaries and medicines	1.1	1.1	1.1	1.1 (1st)
	Increasing Health Insurance (Coverage, Funding, Policy)	1.9	1.9	1.9	1.9 (2nd)
Health Info System	Information availability	1.8	1.6	1.6	1.6 (1st)
	Information use	2.0	2.5	2.1	2.2 (=2nd)
	ICT capacity and equipment including registers and forms for reporting	2.2	1.9	2.3	2.2 (=2nd)
Leadership & Governance	National-level advocacy for maternal health at all levels	3.0	2.4	2.3	2.4 (=1st)
	Supportive supervision visits to health facilities	2.2	2.1	2.5	2.4 (=1st)
	Involvement of health and non-health stakeholders in priority making	3.3	3.0	2.8	2.9 (2nd)
	Issuance of guidelines, charts/algorithms, rules & regulations	2.8	2.9	3.1	3.0 (3rd)
	Decentralization by Devolvement (DbyD)/Basket Fund Vs Local Gov	3.7	4.5	4.3	4.2 (4th)
Number of expert respondents		27	28	155	210

Table 3.6: Mean priority rank of activities/elements within EmOC components, by levels

EmOC Programme Component	Element/Activity	Priority rank of activities within components			
		National	District	Facility	Overall
Human resources	Availability: increased nurse-midwives pre-service students, etc.	1.9	1.9	1.6	1.7 (1st)
	Competence: in service EmOC/IMPACC training	1.6	1.9	1.9	1.8 (2nd)
	Responsiveness & Productivity – provide incentives to staff	2.5	2.2	2.5	2.5 (3rd)
Drugs & Supplies	Procurement of Essential EmOC Drugs	2.0	2.0	2.1	2.1 (1st)
	Procurement (and maintenance) of Equipment & Supplies	2.1	2.6	2.4	2.4 (2nd)
	Availability of emergency transport for patients	2.5	2.3	2.4	2.4 (2nd)
	Improving Integrated Logistics System (ILS)	3.0	3.1	3.1	3.1 (3rd)
Service Provision	Upgrading existing health centres for providing both B&CEmOC functions (building	2.0	2.1	1.8	1.9 (1st)
	Establishing/building maternity waiting homes in hospitals	4.2	3.9	3.5	3.6 (2nd)
	Allocating providers with houses within facilities for emergency	4.9	3.9	3.8	4.0 (3rd)
	Radio call communications & ambulances in selected facilities	3.9	4.4	4.6	4.5 (4th)
	Behavioural Change Communication programme & community mobilization	5.0	6.3	5.5	5.6 (5th)
	Outreach/Mobile services	6.1	6.7	6.2	6.2 (7th)
	Availability of Safe Blood	4.0	5.4	6.6	6.1 (6th)
	Staff meetings	7.3	7.6	7.5	7.5 (9th)
	Waste safe disposal	6.7	7.6	8.0	7.8 (10th)
	Water and Power reliability (source, storage & emergency)	4.8	7.1	7.3	7.0 (8th)
Health Financing	Increased government health budget for salaries and medicines	1.1	1.2	1.1	1.1 (1st)
	Increasing Health Insurance (Coverage, Funding, Policy)	1.9	1.8	1.9	1.9 (2nd)
Health Info System	Information availability	1.5	1.9	1.5	1.6 (1st)
	Information use	2.1	1.9	2.1	2.1 (2nd)
	ICT capacity and equipment including registers and forms for reporting	2.3	2.2	2.5	2.4 (3rd)
Leadership & Governance	National-level advocacy for maternal health at all levels	2.6	2.3	2.3	2.4 (1st)
	Supportive supervision visits to health facilities	2.2	2.4	2.5	2.5 (2nd)
	Issuance of guidelines, charts/algorithms, rules & regulations	2.7	2.5	2.6	2.6 (3rd)
	Involvement of health and non-health stakeholders in priority making	3.1	3.2	3.1	3.1 (4th)
	Decentralization by Devolvement (DbyD)/Basket Fund Vs Local Gov	4.3	4.6	4.6	4.5 (5th)
Number of expert respondents		27	34	153	(n=214)

3.5.5 Priority ranking of EmOC activities

Table 3.6 shows that national-level experts ranked competence of human resources as first priority, whereas district- and facility-level ranked availability of human resources as their first priority for EmOC programme activities. For drugs and supplies activities likewise, national-level experts prioritized procurement (and maintenance) of EmOC equipment and supplies over procurement of essential EmOC drugs, while the latter was a first priority by the district- and facility-level participants. Experts in all levels prioritized upgrading existing health centres to be able to provide both basic and comprehensive emergency obstetric care services. However, while district- and facility-level participants ranked in second place the activity on building maternity waiting homes in hospitals, national-level participants prioritized availability of functioning radio call communications and ambulances in selected facilities as their second activity in priority. Similar to activities for FANC programmes, experts prioritized increased government health budget on salaries and medicines over increased coverage, funding, and policy for health insurance. All experts placed more priority on availability and use of health information than on information, communication, and technology capacity and equipment including registers and forms for reporting. While district- and facility-level participants gave first priority to advocacy of maternal health from the national level to all levels, national-level experts on the other hand maintained their stance for supportive supervision visits to health facilities as first priority on leadership and governance.

3.6 Discussion

This chapter presented the key features required for measuring the implementation strength of a large-scale programme, which are: identifying programme components, developing approaches for measuring components, and devising a scoring mechanism. Study findings in this chapter showed that the six health-system-building blocks were adopted as being the key components for both FANC and EmOC programmes because they constitute the essential functional parts of a large-scale health care programme that can be used to implement programme activities to target populations. An approach was developed using a questionnaire designed to elicit preference weights from maternal health experts. For the weighting scheme,

maternal health experts were used to score the components, and summary scores were generated to represent overall strength of a component for each programme.

Generally, in this study, each respondent's opinion was assumed to carry equal weight regardless of their health-system level or any other personal attribute (such as location (urban/rural) or work experience). Equal weighting of participants' opinions was based on the assumption that those involved in coordinating/training at the national-level, and those involved in planning and supervising at the district-level as well as those involved in day-to-day execution of the FANC and EmOC programme activities at the health facility-level complemented each other's efforts in ensuring the services reached the target populations.

Maternal health experts thought that the human resources component had, on average, the strongest contribution (mean of 24%) to the implementation of FANC programme. This was followed by drugs and supplies (20%), health financing (17%), service provision (17%), health information system (11%) and leadership and governance (11%). For EmOC programme too, human resources component had the strongest contribution to programme implementation (mean of 25%) followed by drugs and supplies (20%). Unlike in FANC implementation, service provision came third (19%) followed by health financing (16%) and health information system and leadership and governance (each 10%). There were slight variations in assigning weights (scoring) among national-, district- and facility-level participants. Participants in one level sometimes scored some components higher than their counterparts in other levels. However, those variations were not significant except for the leadership and governance component of the FANC programme in which district-level and facility-level participants significantly differed. In addition, significant variations were observed on the health financing component for the EmOC programme in which district-level and national-level participants significantly differed.

However, prioritisation of programme activities within each component seemed to vary across the levels – implying that, the needs at the national level were different from the needs at the district level or facility level. For example, for FANC programme activities, district- and facility-level participants thought advocacy of maternal health from the national level to all subsequent levels to be of first priority over conducting supportive supervision visits to health facilities – a first priority by

the national-level experts. Likewise, district- and facility-level experts prioritized procurement of the essential drugs for antenatal care and improving the logistic system for supplying the drugs more important than procuring essential equipment and supplies for antenatal services, which was more important by the national-level experts than by participants in the former level. However, all experts from all three levels ranked the availability of human resources as top priority over and above competence or responsiveness and productivity of the human workers.

In prioritising EmOC programme activities, national-level experts ranked competence of human resources as first priority, whereas district- and facility-level ranked availability of human resources as their first priority. Likewise, for drugs and supplies activities, procurement (and maintenance) of EmOC equipment and supplies was first priority by national-level experts while lower levels prioritised procurement of essential EmOC drugs. Similar observations were seen for the national-level experts who ranked second the building of maternity waiting homes in hospitals whereas district- and facility-level participants had ranked availability of functioning radio call communications and ambulances in selected facilities as their second activity in priority.

For both programmes, upgrading existing health centres to provide both basic and comprehensive functions of emergency obstetric care services, including building operating theatres and labour and delivery wards, were elements given first priority for the service provision component. Increased government health budget for salaries and medicines was the highest-ranked element among those forming the health financing component. For health information system, information availability was ranked higher than information use or ICT capacity and equipment. Similarly, for the leadership and governance component, national-level advocacy for maternal health at all levels and supportive supervision visits to health facilities were the elements given first priority by the maternal health experts.

3.7 Limitations

Results obtained using expert opinion methods have been reported to be "only as good as the judgement of the raters" and that "biases and incomplete knowledge are both possible".[107] To minimize the limitations of expert opinions, a large sample of participants was required. The study recruited 225 individuals and covered all three

key levels of the Tanzanian health system. Previous studies suggest that this number of participants is adequate for reliable statistical analyses.[145 146 147]

It is possible also that use of alternative methods such as discrete choice experiments could have been employed to determine participants' preferences. This would have required creating different scenarios and piloting and improving them for which participants would select their preferred choices. While discrete choice experiments is one approach favourable to health economists, it is possible that its use in identifying programme components for FANC and EmOC programmes could have grouped components in such a way that restricted freedom of choice by the experts as they were free to do so in this study. On the other hand, it is possible that using the 0%-100% scoring method could have potentially introduced heaping of some specific scores among participants. However, piloting of the questionnaire showed that participants were clear in assigning scores adding up to 100% across the six programme components.

3.8 Conclusion

To the best of my knowledge, this study is the first of its kind in Tanzania to collect opinions from maternal health experts on the relative importance of the components in the implementation of FANC and EmOC programmes. The MNPI study involving 55 developing countries (Section 1.4.1) had included Tanzania as one of the study countries and involved between 10 – 25 experts to rate 13 programme attributes with 84 elements. Although the attributes were also based on maternal health, they were not organised in the form of the six building blocks, and neither did the study in Tanzania draw participants from parts of the country other than the main city of Dar es Salaam. In contrast, the current study sampled participants from all three levels of the Tanzanian health system to accommodate opinions from those involved in national-level coordination/training of nurses and midwives, and those involved in planning, supervising and implementing the day-to-day activities of the FANC and EmOC programmes. Future research in this area might include testing the composition of the building blocks in a discrete choice experiment for the sake of finding out how participants across the levels would choose their preferences of the components given different optimal options.

Part III: Results

Chapter 4: Unweighted Scores of the Implementation Strength

4.1 Overview

Using data collected from FANC and EmOC programme activities in study districts, the main objective of this chapter was to analyse the data and generate aggregate scores of implementation strength of both FANC and EmOC programmes in each study district. The scores presented in this chapter are as they were collected from sources and are therefore unweighted. Weighting of these scores is presented in Chapter 5 on weighted composite scores of implementation strength. Section 4.2 describes the methods used in collecting data from study districts related to the implementation strength of FANC and EmOC programmes. Section 4.3 briefly discusses the approach used to analyse the data and Section 4.4 presents the results. Section 4.5 discusses the results and concluding remarks are presented in Section 4.6.

4.2 Methods

4.2.1 Design and study area

The study employed a cross-sectional survey to collect the data on the implementation strength for the FANC and EmOC programmes. The area of study was described in Chapter 2 (Section 2.2) about the 23 nationally representative districts. A brief recount of the study areas is that the 23 districts were selected by the Tanzanian National Bureau of Statistics. The Bureau of Statistics used a two-stage probability proportional to size sampling technique with data from the 2002 Tanzanian National Housing and Population Census. At least one district was selected from each of the seven government administrative zones. Of the 23 districts selected, eight were urban and fifteen were rural.

4.2.2 Data collection

Table 2.3 shows the list of proxy indicators that were used to collect data for each of the health-system-building block (programme component). The final list of programme elements identified for each programme component (Table 3.1) was used as guidance in selecting proxy indicators for each of the six programme components. Selection of indicators took note of the following criteria: if an indicator was popularly used in the literature (such as the number of health workers per 10,000 population), if an indicator was in the list suggested by the World Health Organisation (such as those for the service delivery component), and most importantly, if it was possible to obtain data for the indicator in study districts. The final list of proxy indicators proposed for collecting data for the implementation strength is shown in Table 2.3. The study used six key sources of data for the implementation strength: the District Health Information System (DHIS) central database, the Service Availability and Readiness Assessment (SARA) survey, the Tanzania Health Statistical Tables and Figures 2012 report, district supervision reports, the Mid-Term Expenditure financial reports, and population projections from the Tanzanian National Bureau of Statistics.

The DHIS database contained facility-level data from all health facilities in the 23 study districts. On a regular basis (usually monthly), district coordinators visited all health facilities in their districts and collected summaries of aggregate data. Staff responsible for service provision in each health facility recorded data using standard government HMIS registers before district coordinators' visits. Upon returning to their office stations (usually within the District Medical Officer's premises), FBIS coordinators (described in Section 2.5.2) entered the aggregate data into the DHIS database. The DHIS is a national data warehouse coordinated and centrally managed by the Ministry of Health and Social Welfare with technical assistance from the University of Dar es Salaam's Computer Science Department.[148] Data managers from the Ifakara Health Institute have special privileges to access data for the 23 SPD districts. The author is an employee of the Ifakara Health Institute and was granted permission by the Institute's Director of Research to access DHIS data and to work along with the data managers in extracting and cleaning the data for part of the indicators of this study.

Likewise, the Tanzania Health Statistical Tables and Figures 2012 report was generated by the Ministry of Health and Social Welfare using 2011 data collated from all reporting health facilities in Mainland Tanzania. Under the directive of their respective District Medical Officers, all HMIS focal persons across Mainland Tanzania prepared a comprehensive report using the DHIS data along with other district-level information such as data on inventory, supplies, pharmaceuticals, human resources, and financial services and reported upwards to their respective Regional Medical Officers.[h] Regional Medical Officers ultimately compiled all reports from their districts and submitted the data to the Ministry for further analysis. Analytical reports resulting from the integrated data were prepared and made available for public access. The author obtained an electronic copy of the Health Statistical Tables and Figures 2012 report from the Ministry of Health and Social Welfare during analysis of this study for comparing and complementing data from other sources.

For the SARA survey (described in Section 2.5.3), the author fully participated during preparation of the survey's questionnaires and training tools. Before the SARA survey commenced, data collectors attended a three-day training at a central location. Data collectors consisted of two representatives from each of the 23 study districts, who were usually the HMIS focal person and the FBIS coordinator. During training, participants practiced in groups to familiarize themselves with the use of SARA survey tools. After the training, data collectors visited approximately 68% of the 1,083 health facilities that provided reproductive and child health services in the 23 study districts. To verify the quality and completeness of survey data, senior staff from the Ministry of Health and Social Welfare and those from the Ifakara Health Institute divided themselves among study districts to provide supportive supervision. Separately, the author cleaned the data from the SARA survey and performed the data analysis. The following sections describe the details of the methods used for collecting data for each programme component.

---

[h] HMIS focal persons are staff employed by the District Executive Director to manage and coordinate the district's HMIS data. FBIS Coordinators are staff employed by the Ifakara Health Institute to provide support to and work under the supervision of the District HMIS focal persons.

4.2.2.1 Human resources

The 'number of health workers per 10,000 population' was used as a proxy measure for the human resources component. The World Health Organization established this indicator by setting the number of 23 health workers (doctors, nurses and midwives) per 10,000 population as the minimum number required for delivering 'essential maternal and child health services'.[149] For each study district, numbers of doctors, nurses, midwives and other related health workers were drawn from the Ministry of Health and Social Welfare's official statistics for 2011. Additional data on nurses and midwives specifically trained for providing FANC and basic EmOC services were obtained from Jhpiego (the main organization that provided the training for Mainland Tanzania) in order to further compare study districts on human resources. For the denominator, data on projected population levels (from the census in 2002 projected to 2011) were available from the Tanzanian National Bureau of Statistics. Using the two statistics, ratios of health workers per 10,000 population were generated for each district.

4.2.2.2 Drugs and supplies

Data on availability of tracer drugs and the essential supplies for FANC and EmOC services were part of the data collected by the SARA survey. The survey collected data on availability of syphilis rapid test kits, urine pregnancy test kits, delivery packs, episiotomy scissors, manual vacuum extractor, vacuum aspiration or dilation and curettage kits, neonatal bag and mask, injectable antibiotics, uterotonic and magnesium sulphate. During the SARA survey, respondents in health facilities were asked to respond to YES/NO questions covering all items. Where the response on the availability of items was a 'YES', data collectors verified by observing the items on the survey day.

4.2.2.3 Service delivery

The World Health Organization's handbook on monitoring the building blocks of health systems proposes seven indicators for which to monitor the service delivery component:[133]

- number and distribution of health facilities per 10,000 population;
- number and distribution of inpatient beds per 10,000 population;

- number of outpatient department visits per 10,000 population per year;
- general service readiness score for health facilities;
- proportion of health facilities offering specific services;
- number and distribution of health facilities offering specific services per 10,000 population; and
- specific-service readiness score for health facilities.

The study used three of the seven indicators in consideration of their relevance to FANC and EmOC services and their data availability in study districts. For FANC programme, the three indicators were: the number of health facilities per 10,000 population, the number of inpatient beds per 10,000 population, and the number of first antenatal care visits (instead of number of outpatient visits) per 10,000 population. EmOC programme, the study used only the two indicators of number of health facilities and number of inpatient beds – leaving out the number of ANC visits specifically for FANC. Data for 2011 on the number of health facilities by type for each district, the number of inpatient beds, and the number of first antenatal care visits were obtained from the Ministry of Health and Social Welfare official statistics.

4.2.2.4 Health financing

Districts in Tanzania receive funding from different sources such as basket funding, cost sharing, drug revolving fund, council own sources, insurance schemes, and few other sources. Basket funding and block grants are two main sources of funding for maternal health activities (including FANC and EmOC services). Basket funding includes funding from donors, block grants (for salaries), and funding from the central government. The study accessed published figures on both revenue and expenditure data specific for the 23 study districts. Each year, districts submit their proposed budgets and financial expenditure reports to the Prime Minister's Office for Regional Administration and Local Governments before the parliamentary session for the national budget takes place in June of each year. The Ministry of Regional Administration and Local Governments published in its website[150] the 'Mid-Term Expenditure Framework' financial reports for the two years of 2010/2011 and 2011/2012showing all categories of revenues and expenditure in all districts in Mainland Tanzania (including study districts).

4.2.2.5 Health information system

The study used two proxy indicators for this component: the number of HMIS reports submitted to the District Medical Officers and the timeliness of reporting. Data for the two indicators were available for access from the central DHIS database. A review of the number of reports submitted and their timing was conducted for each district's reports and the number and percentage completion was calculated.

4.2.2.6 Leadership and governance

Quantifying the role played by district health authorities in helping to improve delivery of FANC and EmOC services can be complex. For general oversight, the number of supportive supervision visits conducted each quarter of the year by the Council Health Management Teams was used as the proxy indicator on the leadership and governance component. FBIS coordinators helped reviewing district management records (district supervision reports included) and reported the number of visits Council Health Management Team members conducted in 2010 and 2011.

4.2.3 Data quality assessment and data cleaning

Data is said to be of good quality if it satisfies "the requirements stated in a particular specification and the specification reflects the implied needs of the user."[151] In order to monitor and assess the quality of the data, two data quality control measures were performed: ensuring data collectors had sufficient aptitude of the data type required and appropriate data acquisition skills and ensuring the data sources were reliable. As described in Section 2.5 (training, logistics and data collection), research assistants responsible for collecting implementation strength data from study districts were trained for two days and participated fully in piloting survey tools in one of the 23 study districts – all under the supervision of the author. The author ensured that all nine research assistants had the capability of collecting good quality data from all sources in their respective districts. This was done during training in which research assistants were asked to participate in role-plays - at the end of which, research assistants provided feedback and comments to those performing the plays. During piloting of survey tools and data collection in the pilot district, the author performed spot-checks to data collectors and convened daily

after-work meetings for all five piloting days to discuss challenges and ways to improve data acquisition from sources.

With the exception of the SARA survey, official supervision reports from districts, and the DHIS central database, neither the author nor the data collectors (9 research assistants and 23 FBIS coordinators) were involved in collecting the other types of data (Official statistics from the Ministry of Health and Social Welfare, official statistics on the Mid-Term Expenditure and Revenues from the Prime Minister's Office for Regional Administration and Local Governments, and official population projections for 2010 and 2011 from the Tanzanian National Bureau of Statistics (NBS) of the Ministry of Planning, Economy and Empowerment). However, the author contacted some of the respective staff who provided the official statistics. For example, the author made personal visits and had conversation with senior officials at the Ministry of Health and Social Welfare and at the National Bureau of Statistics. The document containing official statistics from the Ministry of Health and Social Welfare quoted the following in reference to how data was analysed and partly on data quality:

> *"The production of this report involved a number of forums. Initially the district and regional HMIS focal person convened for the data analysis work session under facilitation of central M&E staff. Thereafter data analysis, compilation and interpretation was done by the central M&E and other MOHSW staff including health and public specialists, statisticians, epidemiologist and the region HMIS staff. Through these process (es) regional and national totals, percentage and graphs were prepared. The compiled report was submitted to the MOHSW management for approval and endorsement for publication."*[152]

In addition, the document on population projections for the national housing and population census prepared by the NBS cited its approach on the quality of data as:

> *"Evaluation of the quality of the census data is often necessary to determine whether data collection was properly done and that the data are of acceptable quality. The evaluation and quality control measures were undertaken at pre-enumeration, enumeration and post-enumeration stages. At each stage quality standards were established and maintained to minimize errors in the*

Population and Housing Census undertaking. This resulted in good quality data for most of the basic variables as evaluated in the 2002 Post-Enumeration Survey (PES)."[153]

Regarding financial data from the Prime Minister's Office, there was no independent verification of the quality of data, but the data were considered to be of sufficient quality in faith that all financial data from districts were audited by the Controller and Auditor General before they were released for public access and use. The quality of data from the SARA survey and the DHIS database are as described in Section 4.2.2.

Having gathered all data and before analysis commenced, data for each programme component was assessed individually to ascertain its quality. Quality attributes looked at included completeness (Table 2.7), validity, consistency, and accuracy. For example, although data for the human resources component were available in the DHIS database (which included the number of Assistant Medical Officers, Clinical Officers, Health Officers, Laboratory Officers, Laboratory Technicians, Medical Officers, Nurse Midwives, and Nursing Staff), some of the districts had missing data and some had questionable number of the health workers such as those of Babati and Bagamoyo districts showing as having no medical doctors while the report from the Ministry of Health and Social Welfare showed the two districts had 4 and 6 medical doctors respectively. Data cleaning was only done for data from SARA survey and DHIS database. Two database managers conducted all data cleaning for the two data sources. The author made further verification of the data and contacted database managers upon encountering inconsistency – in which case data managers counterchecked from the database. Follow up resolution was made by data managers by contacting FBIS coordinators of respective districts to further clarify on particular quality issues raised by the author.

4.2.4 Data analysis

Descriptive statistics (percentages, means, and associated standard deviations) were calculated for all indicators of a component. Results were disaggregated by study districts. In order to generate an aggregate score of a component for a particular district, all scores of the indicators constituting that particular component were averaged (arithmetic mean). For example, the component 'Health Information

System' had two indicators: percentage of health facilities reporting specified HMIS reports (reporting rate) and percentage of health facilities reporting within the specified period (reporting timeliness). If, for example, 'Arusha Urban' district had 78% and 84% of its health facilities reporting the required HMIS reports in 2010 and 2011 respectively, and had 46% and 59% of health facilities reporting on time during the two years respectively, then the aggregate score would be the mean of the two scores in the two years. That is, (78% + 84%)/2 = 81% reporting rate, and (46% + 59%)/2 = 52.5% reporting timeliness. The mean aggregate score for Arusha Urban on health information system would therefore be the mean of the two indicators, (81% + 52.5%)/2 = 66.8%.

While other components had more indicators than others, the same approach was used to generate the mean aggregate scores. For indicators that were not expressed in percentages, such as the number of inpatient beds per 10,000 population, an index was calculated to rescale their values between 0% and 100% using the formula $(X_i - X_{min}) \times 100\% / (X_{max} - X_{min})$ where X_i is the value for the ith district (i = 1, 2, 3, …, 23) and X_{min} and X_{max} are the minimum and maximum values of X_i respectively. The approach of rescaling values to between 0% and 100% (also known as 'normalisation') has been used by several global initiatives, including the United Nations Development Programme's Human Development Index, whose indicators of life expectancy, educational attainment and income are combined into a composite index (score).[154] All figures for revenues and expenditure in districts were converted to 'per capita' statistics to take account of the difference in population sizes before normalisation was done.

All districts had complete datasets for their selected indicators except for a few constituent indicators of the health financing component (revenues only) and health information system component (reporting rate and reporting timeliness). Given the level of missingness in the data, no multiple imputation methods were used. According to Acuña and Rodriguez, "rates of less than 1% missing data are generally considered trivial, 1-5% manageable, 5%-15% require sophisticated methods to handle, and more than 15% may severely impact any kind of interpretation".[155] Missingness in data was assessed and found to be less than 5%. For variables/indicators with missing data, simple imputation with means/medians was used in such a way that, if a rural district had missing value for a certain revenue constituent indicator, it was imputed with the value of the median for rural districts,

and if an urban district had missing value for a certain revenue constituent indicator, imputation was done by assigning it with the median value for urban districts. For the two constituent indicators of health information system, three rural districts had missing data and were imputed with the value of the arithmetic mean for rural districts. Summary indicators were tested for any differences between rural districts and urban districts using a two-sample Wilcoxon rank-sum (Mann-Whitney) test. Unless otherwise stated, all P-values reported were those for the Wilcoxon rank-sum test comparing eight urban districts with 15 rural districts.

4.3 Results

4.3.1 Human Resources

Table 4.1 shows the number of health workers, including medical doctors (general and specialists), non-physician clinicians (Assistant Medical Officers, Clinical Officers, and Assistant Clinical Officers), nurses, and midwives for 2011. The table also shows populations for 2011 in each study district. The overall mean number of health workers per 10,000 populations per district was 4 (SD=3). Generally, urban districts were twice well-staffed with health workers (mean of 6 health workers per 10,000 populations) compared to rural districts (mean of 3 core staff per 10,000 populations). The overall mean score for the health resources component was 23% with urban districts scoring more than twice compared to rural districts (39% vs 15%). Arusha Urban, Songea Urban, Bagamoyo, and Ilala Urban were the leading districts with the highest number of health workers per 10,000 population. Districts with the lowest number of health workers per 10,000 population were Uyui, Muleba, Musoma Rural, Kondoa, and Kasulu. Of the health workers available, each district had nurse-midwives specifically trained for providing FANC and basic EmOC services.[i] Table 4.1 shows that urban districts had a mean of 94 FANC-trained nurse-midwives compared to a mean of 55 FANC-trained nurse-midwives in rural districts. In addition, even though the number trained on BEmOC was only a small fraction of the total number of nurse-midwives, urban districts had on average six times more BEmOC-trained nurse-midwives than rural districts (P=0.008).

---

[i] The number of nurse-midwives specifically trained in FANC and EmOC was not included in computing the overall composite score but is presented here for comparison purposes only.

Table 4.1: Programme component indicators on the number of health workers and those specifically trained on FANC and BEmOC

No.	Study District	Population	Medical Doctors (incl Specialists)	Non-Physician Clinicians	Nurses (incl Midwives)	Providers Trained on:		Pharmacists (+ Pharm Techs)	Lab Staff	Health Workers per 10,000 pop	Human Resource Score
						FANC	BEmONC				
1	Arusha Urban	372,719	67	151	333	190	29	16	4	15	100%
2	Ilala	730,108	35	215	226	83	37	12	35	7	38%
3	Iringa Urban	158,592	1	41	45	133	10	2	15	5	30%
4	Kinondoni	1,250,230	24	318	328	123	42	30	25	5	30%
5	Mtwara Urban	126,923	1	13	28	111	16	2	0	3	14%
6	Songea Urban	182,156	8	111	67	54	10	5	10	10	66%
7	Tanga Urban	305,713	10	83	18	0	18	2	4	4	17%
8	Temeke	880,022	15	110	202	57	32	12	8	4	17%
Urban Districts, Mean		500,808	20	130	156	94	24	10	13	6	39%
9	Babati	423,380	4	80	79	67	15	6	6	4	18%
10	Bagamoyo	283,780	6	106	89	60	7	4	9	7	42%
11	Geita	879,836	3	102	175	35	1	8	6	3	13%
12	Kahama	848,728	3	111	107	61	6	2	5	3	9%
13	Kasulu	653,858	3	84	70	58	8	3	6	2	7%
14	Kilosa	600,428	4	177	133	71	4	4	12	5	29%
15	Kondoa	508,353	2	52	33	33	0	2	1	2	2%
16	Mbozi	681,969	3	56	164	31	5	1	1	3	14%
17	Moshi Rural	469,593	6	136	113	141	5	3	5	5	30%
18	Muleba	492,404	7	48	58	0	0	1	2	2	7%
19	Musoma Rural	669,365	2	60	105	39	0	2	2	2	8%
20	Ruangwa	149,634	1	28	37	51	4	1	1	4	22%
21	Singida Rural	497,562	7	71	68	31	0	1	3	3	11%
22	Sumbawanga Rural	499,587	0	80	51	105	6	2	2	3	9%
23	Uyui	362,619	0	24	27	42	3	0	5	1	0%
Rural Districts, Mean		534,740	3	81	87	55	4	3	4	3	15%
Overall, Mean		522,937	9	98	111	69	11	5	7	4	23%
Overall, SD			15	69	90	46	12	7	8	3	22%
2-sample WR-S test for Ho; urban =rural		P-value	0.023	0.208	0.561	0.087	0.0002	0.026	0.092	0.008	0.008

Source: The MOHSW-Health Statistical Tables and Figures for 2011 on data for health workers; FANC and BEmOC training data from Jhpiego Tanzania; WR-S= Wilcoxon Rank Sum; µ = the indicator used only data on Medical Doctors, Non-Physician Clinicians and Nurses as per WHO. In order to transform the indicator health workers per 10,000 population into indices between 0 and 1, the District Score was calculated by dividing the value by 23 (the minimum threshold recommended by WHO).

Table 4.2: Percentage of surveyed health facilities with available Drugs and Supplies on the day of visit

Study District	N ζ = 921	901	854	805	809	799	800	800	1145	917	918	Drugs and Supplies Score		
	Syphilis RT kits	Pregnancy Urine test	Urine PT kits	Delivery pack	Episiotomy scissors	MVE	VA or D&C kits	Neonatal Bag&Mask	Injectable Antibiotics	Injectable Uterotonic	Injectable MgSO4	FANC-Specific	EmOC-Specific	Overall District
Arusha Urban	88%	100%	100%	92%	100%	100%	42%	75%	52%	55%	55%	96%	71%	78%
Ilala	57%	91%	59%	52%	77%	67%	30%	31%	81%	76%	83%	69%	62%	64%
Iringa Urban	79%	96%	96%	91%	91%	82%	36%	27%	85%	67%	67%	90%	68%	74%
Kinondoni	95%	49%	52%	64%	69%	33%	30%	19%	49%	51%	51%	65%	46%	51%
Mtwara Urban	71%	88%	83%	44%	75%	56%	44%	44%	64%	57%	57%	80%	55%	62%
Songea Urban	74%	22%	11%	36%	100%	27%	9%	27%	92%	100%	91%	36%	60%	54%
Tanga Urban	78%	63%	61%	77%	74%	57%	18%	52%	74%	62%	74%	67%	61%	63%
Temeke	78%	100%	98%	37%	83%	53%	26%	16%	70%	75%	83%	92%	55%	65%
Urban, Mean	78%	76%	70%	62	8%	59%	29%	36%	71%	68%	70%	74%	60%	64%
Babati	48%	23%	18%	22%	51%	47%	22%	32%	51%	52%	53%	30%	41%	38%
Bagamoyo	46%	85%	62%	10%	65%	36%	8%	12%	86%	41%	57%	64%	39%	46%
Geita	66%	50%	33%	13%	64%	52%	36%	20%	70%	56%	20%	50%	41%	44%
Kahama	67%	47%	67%	18%	76%	78%	20%	20%	71%	55%	63%	60%	50%	53%
Kasulu	32%	44%	19%	13%	63%	35%	10%	6%	58%	32%	66%	32%	36%	34%
Kilosa	9%	16%	13%	95%	96%	49%	4%	9%	74%	63%	11%	12%	50%	40%
Kondoa	97%	15%	6%	26%	96%	87%	4%	2%	90%	78%	56%	39%	55%	51%
Mbozi	12%	29%	13%	31%	46%	13%	10%	13%	63%	41%	24%	18%	30%	27%
Moshi Rural	78%	63%	58%	16%	100%	95%	16%	13%	88%	93%	92%	66%	64%	65%
Muleba	26%	41%	39%	8%	92%	41%	22%	11%	73%	59%	73%	35%	47%	44%
Musoma Rural	18%	67%	43%	8%	78%	8%	6%	8%	80%	63%	87%	43%	42%	42%
Ruangwa	100%	4%	0%	42%	81%	37%	7%	8%	96%	93%	96%	35%	58%	51%
Singida Rural	75%	100%	100%	56%	100%	45%	18%	16%	84%	81%	98%	92%	62%	70%
Sumbawanga R	21%	30%	17%	84%	70%	9%	5%	5%	60%	25%	28%	22%	36%	32%
Uyui	66%	41%	31%	14%	61%	69%	0%	19%	52%	69%	58%	46%	43%	44%
Rural, Mean	51%	44%	35%	30%	76%	47%	13%	13%	73%	60%	59%	43%	46%	45%
Overall, Mean	60%	55% (31%)	47%	41%	79% (16%)	51%	18%	21%	72%	63%	63%	54%	51%	52%
P-value for Ho	0.028	0.024	0.03	0.01	0.305	0.245	0.005	0.002	0.771	0.478	0.478	0.003	0.0081	0.003

Source: Data from the Service Availability and Readiness Assessment (SARA) survey of the Ifakara Health Institute; RT=Rapid Test, PT=Pregnancy Test, MVE=Manual Vacuum Extractor, VA=Vacuum Aspiration, D&C=Dilation and Curettage, MgSO4=Magnesium Sulphate; N ζ = refers to the total number of facilities that were visited for interviews. For each district, percentages shown reflect the number of facilities visited that provided the services.

4.3.2 Drugs and Supplies

Drugs and supplies for both FANC and EmOC services are shown in Table 4.2. These include syphilis rapid test kits, urine and test kits for pregnancy, delivery packs, episiotomy scissors manual vacuum extractors, vacuum aspiration or dilation and curettage kits, neonatal bags and masks, injectable antibiotics, injectable uterotonic and injectable magnesium sulphate. Such drugs and supplies were on average available in 52% (SD=14) of all visited health facilities in the study districts. Episiotomy scissors and injectable antibiotics were on average the most stocked in health facilities (mean facility availability of 79% and 73% respectively) followed by injectable uterotonic and injectable magnesium sulphate, each of which were available in 63% of surveyed facilities. Health facilities were least likely to stock kits for vacuum aspiration or dilation and curettage, or neonatal bags and masks, which were stocked in only 18% and 21% of surveyed health facilities respectively.

Health facilities in urban districts were more likely to stock drugs and supplies than health facilities in rural districts, with mean availability of 64% and 45% respectively. There was a significant difference between rural and urban districts in stocking of drugs and supplies (P=0.003). On average, Arusha Urban, Iringa Urban, and Singida Rural had the highest scores in overall stocking of FANC and EmOC drugs and supplies, with, respectively, 78%, 74%, and 70% of health facilities with stocks on survey day. Mbozi, Sumbawanga Rural, and Kasulu districts had the lowest availability of drugs and supplies for FANC and EmOC care (27%, 32%, and 34% respectively).

4.3.3 Service delivery

Table 4.3 shows district achievement levels for the service delivery component (on both FANC and EmOC services). For each 10,000 population, study districts had an overall mean (and standard deviation) of 1.5 (0.5) health facilities, 10 (9) hospital beds and 411 (123) first attending pregnant women in antenatal care clinics. For FANC service delivery, study districts had an overall mean (and standard deviation) of 44% (17%) with Arusha Urban, Ilala, Iringa Urban, and Mtwara Urban having the top scores compared to other remaining districts. Districts with the lowest FANC service delivery scores were Musoma Rural, Muleba, and Uyui. For EmOC service delivery, study districts had an overall mean (and standard deviation) of 40% (21%) with Iringa Urban, Arusha Urban, Mtwara Urban and Ilala leading other districts in having to scores. Districts with the lowest EmOC service delivery scores were Geita, Musoma Rural, and Uyui. In general, urban districts had a higher number of health facilities per 10,000 population

Table 4.3: Implementation strength indicators for the service delivery component

Study district	Pop 2011	No of HFs	HFs per 10,000 pop		Beds per 10,000 pop		No 1st ANC visits	1st ANC visits per 10,000 pop		FANC SD Score	EmOC SD Score
			No.	%	No.	%		No.	%		
Arusha Urban	372,719	66	1.8	52%	25	87%	21,707	582	93%	77%	73%
Ilala	730,108	199	2.7	100%	9	30%	43,588	597	96%	75%	70%
Iringa Urban	158,592	26	1.6	45%	28	98%	6,638	419	59%	67%	69%
Kinondoni	1,250,230	242	1.9	60%	9	29%	51,072	409	57%	49%	47%
Mtwara Urban	126,923	24	1.9	58%	23	79%	4,105	323	39%	59%	63%
Songea Urban	182,156	24	1.3	29%	29	100%	6,238	342	43%	57%	61%
Tanga Urban	305,713	55	1.8	53%	21	72%	8,171	267	27%	51%	57%
Temeke	880,022	135	1.5	40%	7	24%	33,310	379	50%	38%	35%
Urban, Mean	500,808	97	1.9	60%	13	44%	21,854	436	62%	55%	59%
Babati	423,380	54	1.3	27%	23	78%	15,462	365	48%	51%	52%
Bagamoyo	283,780	65	2.3	78%	6	21%	12,400	437	63%	54%	52%
Geita	879,836	67	0.8	1%	4	14%	54,137	615	100%	38%	23%
Kahama	848,728	63	0.7	0%	6	21%	38,703	456	67%	29%	20%
Kasulu	653,858	87	1.3	30%	12	41%	33,471	512	78%	50%	42%
Kilosa	600,428	80	1.3	30%	8	26%	21,275	354	45%	33%	31%
Kondoa	508,353	74	1.5	36%	6	19%	16,456	324	39%	31%	29%
Mbozi	681,969	94	1.4	32%	7	25%	29,502	433	62%	39%	34%
Moshi Rural	469,593	81	1.7	50%	16	56%	8,121	173	7%	37%	45%
Muleba	492,404	43	0.9	7%	13	44%	6,838	139	0%	17%	21%
Musoma Rural	669,365	65	1.0	12%	2	4%	19,214	287	31%	16%	12%
Ruangwa	149,634	25	1.7	47%	6	20%	3,881	259	25%	31%	32%
Singida Rural	497,562	66	1.3	29%	9	31%	19,231	387	52%	37%	34%
Sumbawanga	499,587	118	2.4	82%	4	12%	24,122	483	72%	55%	51%
Uyui	362,619	40	1.1	18%	0	0%	16,466	454	66%	28%	19%
Rural, Mean	534,740	68	1.3	27%	8	26%	21,285	398	54%	36%	33%
Overall, Mean	522,937	78	1.5	38%	10	32%	21,483	411	57%	44%	42%
Overall, SD		53	0.5	25%	9	30%	14,918	123	26%	17%	18%
test Ho: urban=rural		P-value		0.024		0.007			0.847	0.002	0.008

SD=Service Delivery; HFs=Health Facilities; ANC=Antenatal Care; Overall District Score is the mean of Number of Health Facilities per 10,000 population and Number of beds per 10,000 population scores and the number of first ANC visits per 10,000 population

compared to rural districts, with a mean of 1.9 versus 1.3. Ilala district had the highest number of 2.7 facilities per 10,000 population. Likewise, for the number of hospital beds and first antenatal care visits per 10,000 populations, urban districts had better figures than rural districts, with a mean of 13 hospital beds per 10,000 population (versus 8 in rural districts, P=0.007) and 436 first antenatal care visits per 10,000 population (versus 398 in rural districts, P=0.847).

4.3.4 Health financing

Results on health financing in study districts are divided into two groups: district revenue for general activities, shown in Table 4.4, and district expenditure on health, shown in Table 4.5. All currencies are expressed in Tanzanian Shillings (TZS, one United Kingdom Pound Sterling was equivalent to TZS 2,267 and one United States Dollar was equivalent to TZS 1,468 on 31st December 2010).[156] The main sources of revenues in districts are user charges, insurance funds (both the National Health Insurance Fund and Community Health Fund, and other insurance funding mechanisms), and basket funds. Basket funding includes funding from donors, block grants (for salaries), and funding from the central government. Other sources include district-generated funds and other minor funding mechanisms within districts and from vertical programmes that are not discussed here.

District revenues are directed for expenditure mainly on the following sectors: education, health, agriculture, district roads, water, local administration, and other development activities. However, user fees, insurance, MMAM funds, and sizeable amount of basket funding are earmarked for health expenditure. The overall median annual per capita total revenue in study districts for 2011 was TZS 11,745 – with urban districts having higher figures (TZS 16,739) than rural districts (TZS 9,560), the difference which was significant (P=0.0201). After 'normalizing' the per capita figures into the scale between 0 and 1, Tanga Urban scored 100%, having the highest figure of per capita revenue among the 23 districts. Musoma Rural scored 0%, having the lowest per capita revenues. The overall district score on revenue for all districts was 37%.

Expenditure data on health were grouped into two broad categories of recurrent expenditure and development-based expenditure. Recurrent expenditure included spending on personal emoluments and other recurring charges whereas development-based expenditures were on expenditure directed to planned activities aimed for development in the health sector. For the

Table 4.4: Overall District Revenues in Tanzanian Shillings for 2011

Study district	Cost Sharing	Insurance Funds	Basket Funds	Block Grants		Development Funds	District's Own Funds	MMAM	Global Fund	per capita revenue	District Score
				Other Charges	Personal Emoluments						
Arusha Urban	63,100,000	33,568,799	509,737,800	136,307,025	3,048,344,228	3,876,473,062	82,634,000	76,006,000	155,978,077	21,194	87%
Ilala	1,328,424,472	24,900,000	1,235,826,875	315,647,400	5,532,341,900	205,000,000	197,000,000	76,021,000	364,839,656	12,710	42%
Iringa Urban	47,000,000	12,000,000	234,463,300	151,851,000	1,210,207,980	1,523,358,520	50,548,000	58,952,000	5,054,800	20,767	84%
Kinondoni	1,300,000,000	24,900,000	1,889,068,400	754,438,000	7,426,811,800	1,658,061,000	1,457,360,0	76,021,000	234,439,000	11,855	38%
Mtwara Urban	8,059,000	10,400,000	169,811,200	115,070,000	612,253,478	1,523,358,520	82,634,000	76,036,000	228,174,242	22,264	92%
Songea Urban	28,740,150	30,068,500	262,451,950	84,057,522	981,002,649	108,482,000	14,769,976	108,482,00	13,670,000	8,958	23%
Tanga Urban	100,900,000	24,900,000	4,368,059,251	447,783,300	195,991,996	1,523,358,520	114,720,000	76,021,000	456,678,660	23,906	100%
Temeke	828,345,399	24,900,000	1,597,999,400	267,379,300	5,467,150,838	1,523,358,520	45,000,000	76,021,000	505,407,084	11,745	37%
Urban, Median	82,000,000	24,900,000	872,782,338	209,615,150	2,129,276,104	1,523,358,520	82,634,000	76,021,000	231,306,621	16,739	63%
Babati	37,734,242	63,087,000	727,817,625	335,793,000	2,882,528,455	143,000,000	26,400,000	199,607,00	204,081,881	10,912	33%
Bagamoyo	37,155,442	34,916,000	506,586,000	178,817,000	1,412,398,684	120,000,000	15,000,000	314,464,00	93,497,579	9,560	26%
Geita	50,000,000	234,000,000	1,704,151,600	628,566,000	3,392,506,708	30,000,000	30,000,000	428,265,00	1,687,788,253	9,303	25%
Kahama	116,553,045	56,338,985	702,540,500	369,492,000	2,072,195,796	620,986,419	5,000,000	314,464,00	141,585,202	5,183	4%
Kasulu	51,836,800	40,882,833	1,223,352,850	355,248,600	2,029,691,021	143,000,000	11,250,000	435,268,00	172,833,542	6,826	12%
Kilosa	43,323,000	143,903,000	1,205,620,400	423,989,000	2,534,515,392	229,385,000	5,500,000	314,464,00	4,389,316,492	15,472	57%
Kondoa	177,312,400	65,032,571	1,023,963,100	442,841,000	2,166,012,650	360,000,000	50,000,000	347,000,00	86,321,169	9,282	25%
Mbozi	41,661,500	116,023,469	1,003,308,325	519,000,000	2,539,339,915	143,000,000	17,500,000	314,464,00	172,833,542	7,137	14%
Moshi Rural	154,020,000	64,673,716	798,000,000	2,326,347,625	3,128,807,524	143,000,000	10,000,000	388,234,00	172,833,542	15,302	56%
Muleba	40,000,000	46,000,000	883,730,000	549,949,548	4,364,851,608	529,089,000	11,250,000	222,089,00	105,706,000	13,714	48%
Musoma Rural	15,000,000	23,225,000	688,796,000	294,916,000	1,423,324,065	30,000,000	15,000,000	314,464,00	172,833,542	4,448	0%
Ruangwa	33,538,000	22,359,818	324,456,900	433,415,000	752,888,086	143,000,000	10,500,000	81,538,000	30,322,697	12,243	40%
Singida Rural	41,661,500	112,535,913	973,843,809	342,129,760	1,163,213,000	52,000,000	5,000,000	314,464,00	1,885,912,972	9,829	28%
Sumbawanga R	41,661,500	12,812,000	888,676,650	512,067,378	1,107,910,629	143,000,000	10,000,000	314,464,00	172,833,542	6,412	10%
Uyui	33,000,000	46,000,000	812,487,222	1,654,053,741	716,674,200	143,000,000	11,250,000	281,928,00	2,384,407,901	16,775	63%
Rural, Median	41,661,500	56,338,985	883,730,000	433,415,000	2,072,195,796	143,000,000	11,250,000	314,464,00	172,833,542	9,560	26%
Overall, Med	43,323,000	34,916,000	883,730,000	369,492,000	2,072,195,796	143,000,000	15,000,000	281,928,00	172,833,542	11,745	37%
Overall, IQR	89,668,702	40,282,180	615,219,725	243,003,370	1,953,014,062	960,810,970	37,250,000	277,146,00	396,675,294	6,268	32%
2-sample Wilcoxon Rank Sum test for Ho; urban =rural					P-value					0.01	

Source: Official Statistics from the Ministry of Health and Social Welfare for 2011. MMAM= Swahili acronym for Primary Health Services Development Programme

Table 4.5: Overall District Expenditure in Thousands of Tanzanian Shillings

No.	Study District	2010 Population	2010 Population	District Health Spending 2010/2011		District Health Spending 2011/2012		District Health Spending per Capita[μ]		2-yr Health Spending per capita	District Score
				Recur+Dev	% of Total	Recur+Dev	% of Total	2010/2011	2011/2012		
1	Arusha Urban	362,484	372,719	6,747,449.0	23.4	8,069,967.7	16.7	18.6	21.7	20.2	100%
2	Ilala	722,027	730,108	8,787,022.7	15.4	14,512,019.7	14.3	12.2	19.9	16.0	75%
3	Iringa Urban	152,649	158,592	1,536,925.8	10.5	2,083,803.3	11.0	10.1	13.1	11.6	47%
4	Kinondoni	1,237,145	1,250,230	10,144,602.8	15.1	16,594,593.6	19.6	8.2	13.3	10.8	42%
5	Mtwara Urban	122,588	126,923	693,469.7	4.7	1,207,281.2	5.0	5.7	9.5	7.6	22%
6	Songea Urban	175,658	182,156	1,343,386.1	8.9	1,884,901.1	10.5	7.6	10.3	9.0	31%
7	Tanga Urban	298,881	305,713	2,830,892.0	9.0	3,414,093.1	7.6	9.5	11.2	10.3	39%
8	Temeke	870,999	880,022	11,040,277.0	16.7	15,990,884.7	23.3	12.7	18.2	15.4	71%
Urban Districts, Median		330,683	339,216	4,789,170.5	13	5,742,030.4	13.0	14.5	16.9	15.7	73%
9	Babati	407,780	423,380	3,444,565.2	17.2	2,743,412.4	12.9	8.4	6.5	7.4	21%
10	Bagamoyo	277,673	283,780	2,982,592.2	12.5	3,800,368.9	11.3	10.7	13.4	12.1	50%
11	Geita	856,074	879,836	1,203,192.1	3.0	5,752,816.0	13.4	1.4	6.5	4.0	0%
12	Kahama	815,177	848,728	4,926,712.2	15.7	5,950,618.5	17.2	6.0	7.0	6.5	16%
13	Kasulu	631,318	653,858	3,491,676.0	14.1	4,747,929.7	12.9	5.5	7.3	6.4	15%
14	Kilosa	587,967	600,428	5,732,062.6	19.1	5,095,207.8	16.1	9.7	8.5	9.1	32%
15	Kondoa	499,407	508,353	3,658,520.3	12.5	3,723,330.1	11.5	7.3	7.3	7.3	21%
16	Mbozi	662,287	681,969	4,361,231.4	14.0	3,909,763.5	10.9	6.6	5.7	6.2	13%
17	Moshi Rural	462,085	469,593	3,726,778.0	9.5	4,328,314.4	9.1	8.1	9.2	8.6	29%
18	Muleba	475,742	492,404	1,954,165.5	8.6	3,179,853.4	11.2	4.1	6.5	5.3	8%
19	Musoma Rural	419,962	669,365	3,163,098.0	13.4	2,252,158.2	8.2	7.5	3.4	5.0	6%
20	Ruangwa	146,990	149,634	1,336,807.7	13.0	1,684,941.3	15.5	9.1	11.3	10.2	38%
21	Singida Rural	486,901	497,562	3,096,756.0	14.6	3,299,419.7	13.9	6.4	6.6	6.5	15%
22	Sumbawanga Rural	482,987	499,587	2,490,312.4	15.1	3,187,694.4	13.7	5.2	6.4	5.8	11%
23	Uyui	352,560	362,619	2,769,195.5	15.6	2,393,527.7	13.1	7.9	6.6	7.2	20%
Rural Districts, Median		482,987	499,587	3,163,098.0	14.0	3,723,330.1	12.9	6.5	7.5	7.0	16%
Overall, Median		475,742	497,562	3,163,098.0	14.0	3,723,330.1	12.9	6.6	7.5	7.1	22%
Overall, IQR		321,082	341,501	2,421,732.9	5.5	2,855,541.9	3.9	3.4	5.6	4.1	25%
2-sample WR-S test for Ho; urban =rural				P-value						0.0012	0.0012

Source: Financial data from Prime Minister's Office for Regional Administration and Local Governments

μ = Total yearly health spending divided by number of population estimates in that year; Recur=Recurrent expenditure on personal emoluments and other charges; Dev= Expenditure on development activities

two-year financial data, the overall median health expenditure per capita in study districts was TZS 7,075. Notably, urban districts had over twice the per capita health spending of rural districts, with TZS 15,721 compared to TZS 7,009. The observed difference in per capita spending on health was significant (P=0.0012). While urban districts had increased their median per capita health spending from TZS 14,483 in 2010 to TZS 16,927 in 2011 (about 17% increase), rural districts had slightly less increase in health spending per capita, from TZS 6,549 in 2010 to TZS 7,453 in 2011 (14% increase). On overall district performance, Arusha Urban had 100% score followed by Ilala and Temeke districts with 75% and 71% respectively. Geita had the lowest score with 0%. Musoma Rural and Muleba districts performed poorly in overall health spending per capita, with scores of 6% and 8% respectively.

4.3.5 Health information system

Overall reporting of HMIS data from health facilities to respective district medical offices was better in 2010 than it was in 2011. As Table 4.6 shows, the mean percentage of reports submitted by health facilities to districts dropped from 49% (SD=16%) in 2010 to 38% (SD=22%) in 2011. In 2010, urban districts had slightly similar HMIS report submission (mean reporting rate of 51%), compared to rural districts (mean reporting rate of 50%). however, in 2011, urban districts had better reporting rate (with a mean of 38%) than rural districts (mean of to 34%). Overall timeliness of HMIS reports from health facilities improved in 2011 compared to 2010 with means (SD) of 21% (15%) in 2010 and 39% (51%). Combining the two indicators, districts had an overall score of 32% (SD=10%), with nearly similar scores between urban districts (33%) and rural districts (32%). Geita, Arusha Urban, and Temeke had the top scores on health information system, with 55%, 46% and 45% respectively, while Sumbawanga Rural (19%), Kinondoni (19%), Moshi Rural (22%), and Kasulu (22%) had the lowest scores. Differences between urban and rural districts did not have significant test results.

Table 4.6: Implementation strength indicators for the Health Information System component

No.	Study District	2010					2011							District Score
		Actual Reports	Expected Reports	Percent Reporting	Reports On Time	Percent On Time	Actual Reports	Expected Reports	Percent Reporting	Reports On Time	Percent On Time	Percent Reporting	Percent On time	
1	Arusha Urban	1,431	2,430	59%	715	50%	1,272	2,430	52%	285	22%	56%	37%	46%
2	Ilala	3,606	7,219	50%	341	9%	3,247	7,219	45%	184	6%	47%	8%	28%
3	Iringa Urban	979	1,901	51%	348	36%	872	1,901	46%	256	29%	49%	33%	41%
4	Kinondoni	4,458	9,528	47%	283	6%	1,732	9,528	18%	1	0%	32%	5%	19%
5	Mtwara Urban	938	1,590	59%	230	25%	760	1,590	48%	52	7%	53%	17%	35%
6	Songea Urban	600	1,926	31%	32	5%	1,041	1,926	54%	10	1%	43%	3%	23%
7	Tanga Urban	2,173	4,127	53%	17	1%	931	4,127	23%	314	34%	38%	11%	24%
8	Temeke	4,651	8,341	56%	2,053	44%	4,220	8,341	51%	1,284	30%	53%	38%	45%
Urban Districts, Mean		2,355	4,633	51%	502	22%	1,759	4,633	42%	298	16%	46%	19%	33%
9	Babati	1,837	3,803	48%	779	42%	2,114	3,803	56%	39	2%	52%	21%	36%
10	Bagamoyo	4,503	10,145	44%	2,000	44%	3,477	10,145	34%	1,034	30%	39%	38%	39%
11	Geita	3,585	5,052	71%	831	23%	3,963	5,052	78%	1,762	44%	75%	34%	55%
12	Kahama	3,731	5,088	73%	796	21%	3,208	5,088	63%	1,125	35%	68%	28%	48%
13	Kasulu	1,213	5,591	22%	37	3%	3,470	5,591	62%	63	2%	42%	2%	22%
14	Kilosa	3,716	5,769	64%	546	15%	322	5,769	6%	231	72%	35%	19%	27%
15	Kondoa	3,109	4,988	62%	865	28%	1,691	4,988	34%	504	30%	48%	29%	38%
16	Mbozi	1,085	4,389	25%	80	7%	229	4,389	5%	470	205%	15%	42%	28%
17	Moshi Rural	1,761	5,741	31%	69	4%	328	5,741	6%	470	143%	18%	26%	22%
18	Muleba	2,056	3,695	56%	468	23%	2,433	3,695	66%	97	4%	61%	13%	37%
19	Musoma	1,692	5,221	32%	305	18%	253	5,221	5%	78	31%	19%	20%	19%
20	Ruangwa	2,538	2,917	87%	516	20%	387	2,917	13%	470	121%	50%	34%	42%
21	Singida Rural	2,028	4,676	43%	139	7%	1,734	4,676	37%	459	26%	40%	16%	28%
22	Sumbawanga Rural	3,542	7,580	47%	204	6%	1,244	7,580	16%	126	10%	32%	7%	19%
23	Uyui	1,680	3,621	46%	99	6%	1,718	3,621	47%	122	7%	47%	7%	27%
Rural Districts, Mean		2,538	5,218	50%	516	18%	1,771	5,218	35%	470	51%	43%	22%	32%
Overall Districts, Mean		2,474	5,015	49%	511	21%	1,767	5,015	38%	410	39%	44%	21%	32%
Overall Districts, SD			2,327	16%	553	15%	1,275	2,327	22%	464	51%	15%	13%	10%
2-sample WR-S test for Ho: urban =rural		P-value		0.302		0.519			0.519		0.897	0.439	0.519	1.000

HMIS = Health Management Information System

4.3.6 Leadership and governance

As a proxy indicator for leadership and governance, supportive supervision visits are an intervention that district health authorities use to improve provision of services in health facilities. Table 4.7 shows the number of planned and achieved visits by study district for 2010 and 2011. Each district has various ways in which they plan and achieve their supportive supervision. Some districts plan quarterly visits (that is, four visits per year) and some districts plan monthly visits (twelve visits per year) during which, supervisors visit all health facilities in their respective district.

Table 4.7: Implementation strength on the Leadership and Governance component

No.	Study District	Number of Supportive Supervision Visits						District Score
		2010			2011			
		Planned	Achieved	% achieved	Planned	Achieved	% achieved	
1	Arusha Urban	216	156	72%	240	192	80%	76%
2	Ilala	398	282	71%	458	299	65%	68%
3	Iringa Urban	12	12	100%	12	7	58%	79%
4	Kinondoni	204	100	49%	204	120	59%	54%
5	Mtwara Urban	4	3	75%	4	4	100%	88%
6	Songea Urban	24	18	75%	24	24	100%	88%
7	Tanga Urban	216	106	49%	216	118	55%	52%
8	Temeke	1452	1016	70%	1284	1027	80%	75%
	Urban Districts			70%			75%	73%
9	Babati	12	11	92%	12	12	100%	96%
10	Bagamoyo	44	44	100%	44	44	100%	100%
11	Geita	222	222	100%	276	260	94%	97%
12	Kahama	156	95	61%	156	54	35%	48%
13	Kasulu	144	79	55%	144	55	38%	47%
14	Kilosa	12	7	58%	12	9	75%	67%
15	Kondoa	12	5	42%	12	10	83%	63%
16	Mbozi	12	10	83%	12	9	75%	79%
17	Moshi Rural	480	304	63%	480	413	86%	75%
18	Muleba	204	162	79%	204	160	78%	79%
19	Musoma Rural	60	40	67%	75	68	91%	79%
20	Ruangwa	12	9	75%	12	10	83%	79%
21	Singida Rural	240	116	48%	240	180	75%	62%
22	Sumbawanga Rural	12	8	67%	12	7	58%	63%
23	Uyui	156	102	65%	156	71	46%	56%
	Rural Districts			70%			75%	73%
	Overall Mean			71%			76%	74%
	Overall SD			70%			77%	75%
	2-sample Wilcoxon Rank Sum test: urban =rural		P-value	0.796			0.923	0.897

However supportive supervision visits were planned, districts reported the number of visits they achieved as a proportion to the number planned. Overall, the mean percent supportive supervision visits achieved by districts increased by 5% in 2011 from 70% in 2010 for both rural and urban districts. On average, both rural and urban districts had similar scores on planned supportive supervision visits. However, Bagamoyo (100%), Geita (97%), and Babati (96%) (all rural districts), had the topmost scores on supportive supervision visits for both 2010 and 2011. The lowest-achieving districts were Kahama, Kasulu, and Tanga Urban, with an average achievement across 2010 and 2011 of 47%, 48%, and 52% respectively.

4.4 Discussion

This chapter presented the results of programme implementation strength data from study districts for all six programme components. A general discussion of where districts sit overall in relation to six dimensions is presented in Chapter 5 on composite scores which incorporated the weighting scales generated in Chapter 3.

Human resources: Results presented in this chapter indicate that there was much lower staffing of health workers in study districts than the number recommended by the World Health Organization. Study districts were reported to have an average of 4 health workers (physicians, nurses and midwives) per 10,000 population as opposed to the recommended number of 23 per 10,000 population.[149] Even though the study did not research on the effect of low staffing on FANC and EmOC services, the Global Health Workforce Alliance reported that shortage of human resources for health is a threatening limitation to the achievement of the health-related MDGs (MDG 4, 5 and 6).[157] Findings by this study on the inequitable distribution between urban and rural districts agree with part of findings by Kieffer and Kiragu.[32] Kieffer and Kiragu found that lower staffing levels for most of the sub-Saharan African countries was frequently affected by the inequitable distribution between urban and rural areas, insufficient training, weak management, and low staffing ratios.

Service delivery: The service delivery component was evaluated based on the number per 10,000 population of: health facilities, inpatient beds, and number of first antenatal care visits. Urban districts were reported to have better scores on all three indicators of service delivery than rural districts. The study reported an overall average of 1.5 health facilities per 10,000 population (equivalent to 75 facilities per

500,000 population). This proportion of health facilities seems to be over and above the proportion recommended by the World Health Organization of 5 facilities per 500,000 population for provision of EmOC. While the proportion of health facilities per 500,000 population seems overwhelming (25 times more than WHO recommendations), the number of inpatient beds in health facilities especially those available for use by pregnant and postpartum women needs to be increased. This fact was revealed in the City of Dar es Salaam where there has been constant overcrowding of the labour wards to the extent that up to three pregnant women were reported to sharing one bed.[158] [The City of Dar es Salaam is formed of three municipal districts of Ilala, Temeke and Kinondoni – all of which were among the 23 districts of this study]. Overcrowding can heavily influence provision of essential services to women in labour such as bed-side visiting of family members for emotional support, privacy and dignity of women, and provision of other obstetric services. All of such important services to women in labour can also have a contributing effect to the coverage of services in study districts. [159] [160]

With about 94% and 96% of pregnant women attending first antenatal care visits in study districts and in the national survey respectively,(TDHS 2010) Tanzania appears to have relatively high coverage of first antenatal care visits compared to most countries in sub-Saharan Africa.[161] While high coverage of first antenatal care visits is an important indicator of the service delivery component and helps pregnant women in receiving initial counselling services that are important during pregnancy and overall health during pregnancy, neither the indicator nor the indicator on at least four antenatal care coverage (WHO recommended) is likely to guarantee that the woman would have received the required interventions.[162] The indicators simply assume that, compared to pregnant women with less visits or no visit at all, pregnant women attending at least four antenatal care visits have increased likelihood of receiving the key pregnancy-related interventions. The study used only three of the seven WHO-recommended indicators for the service delivery component. It is very possible that by including all seven indicators, results on the aggregate scores would have shown a different picture in the study districts.

Drugs and supplies: The study has reported that around half of health facilities stocked the essential drugs and supplies for FANC and EmOC services - with urban districts having better stocking than rural districts. Drugs and supplies in the Tanzanian health system are managed through the Integrated Logistic System

(ILS). Figure 4.1 shows how the ordering and flow of supplies under ILS is managed. The manual with procedures on how to roll-out the ILS in Tanzania points out that all health facilities in the country are responsible for managing their funds for purchasing of drugs and supplies, with funds being allocated from the central government.[163]

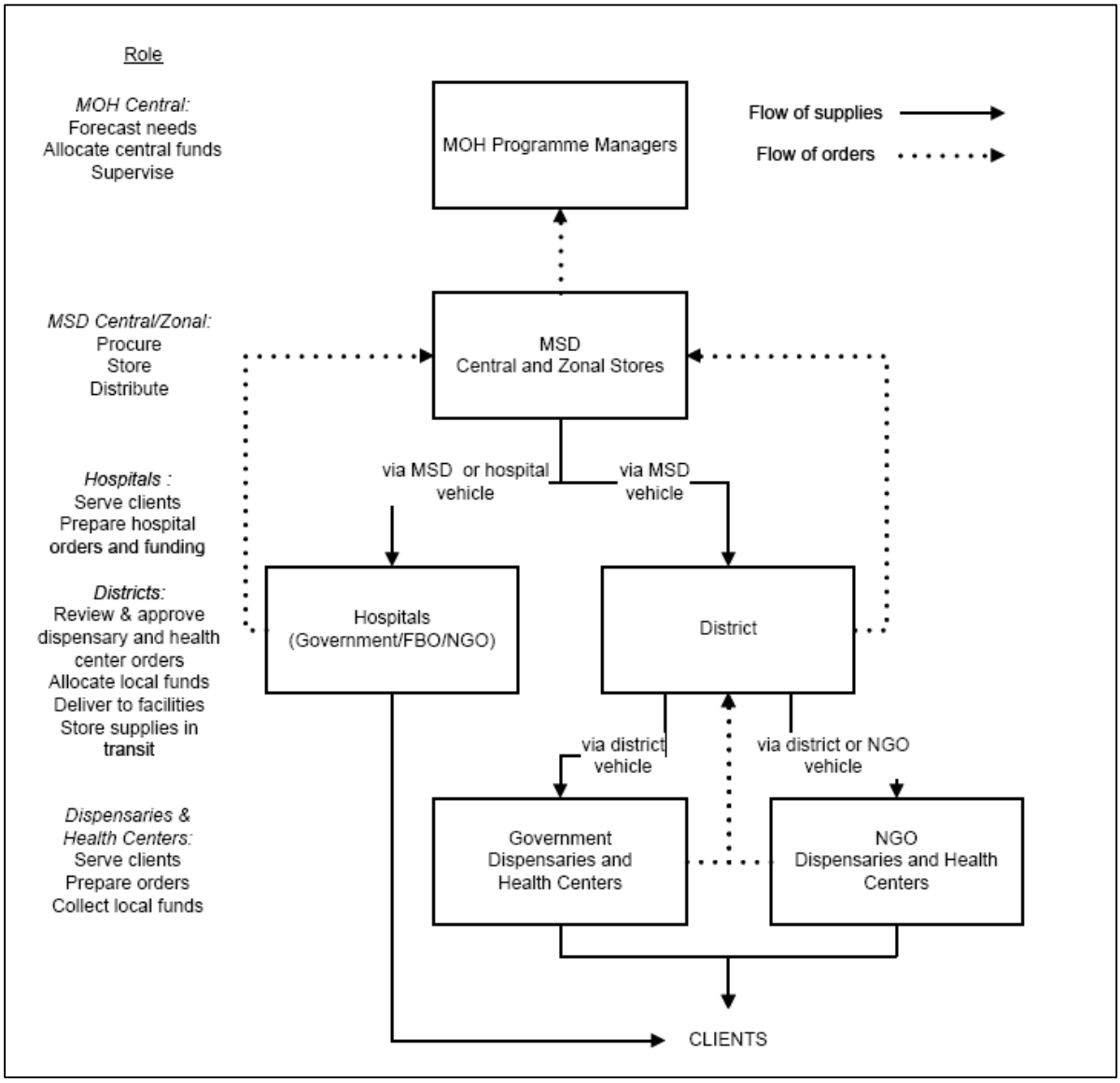


Figure 4.1: Ordering and flow of supplies under the Integrated Logistics System
Adopted from Integrated Logistics System Procedures Manual Roll-Out Version.[163]

While all planning and management of the ILS seems well in order, many areas especially those in rural areas still experience lengthy stock outs and it is possible that there is still a need for supporting health facilities in correct ordering according to their needs as well as in ensuring stocks reach their destinations in time. Poor stocking of the essential drugs and supplies in health facilities can be compounded

by several challenges. For example, in 2010, a study by the Deliver Project in Tanzania assessed over 1,000 facilities across the country on how ordering of drugs was processed.[164] The study reported four key challenges: districts not submitting the appropriate forms 'Request and Requisition' (R&R) on time as required, some districts submitting the forms with "incomplete and missing critical information such as facility name, facility code, and date of submission", other districts did not verify forms from health facilities for completeness or accuracy before forwarding them to the Medical Stores Department (MSD), and lengthy delays "between the completion of forms at the facility and their arrival at MSD". It is possible that districts in this study especially rural districts were most affected by one or more of the four challenges identified by the Deliver Project study.

Health financing: Districts collect revenues from various sources – the main sources being user charges, insurance funds, and basket funds. Basket funding is the biggest contributor of district revenues and includes funding from donors, block grants (for salaries), and funding from the central government. Expenditure on health is mainly financed through revenues collected from user fees, health insurance, and some of the basket funding. There are annual fluctuations on district expenditure for health depending on priority areas set by counsellors within each district. However, district revenues are allocated to several areas that are generally priority areas and include education, health, agriculture, infrastructure, water and sanitation, local administration, and other development activities.

While services for pregnant women and children under the age of five years in the country's public health facilities are supposed to be fee-exempt, user fees are still the norm at all levels of the health system.[165] As a result, women using FANC and EmOC services especially those in the most disadvantaged socioeconomic status (lowest quintile) are most at risk of not accessing the services. Besides the universal health coverage getting global attention, financial protection of people outside the formal employment sector still attracts more debate on how to address equity implications through different financing mechanisms.[166] Some studies have suggested that funding of health services should move away from out-of-pocket payments to prepayment funding mechanisms, which will help protecting people in lower income groups and eventually reduce inequities.[167] With low health insurance coverage to populations and unaffordability of health services, those not covered by health insurance (for example the most vulnerable and the poor) are unlikely to have a

guaranteed access to health services.[168] Even though the Tanzanian government signed the Abuja Declaration in 2001 (which requires governments to commit to spending 15% of the total government budget on health),[169] the World Health Organisation reported that the country has not lived up to its commitment to the Abuja declaration with its health spending falling from 17% of total budget in 2006 to only 11% in 2011.[170] Despite the government's strategy for providing sustainable and integrated health financing, the country's health financing system is still fragmented. Even though there are various sources of funds and funding programmes in Tanzania (for example the comprehensive national health insurance fund (NHIF) and the voluntary CHF), still many people continue facing financial hurdles from accessing the services.[171]

Health information system: Results by this study have indicated that there was sufficient reporting of HMIS reports from health facilities to district offices and that urban districts had better HMIS data reporting to districts than rural districts (with a rate of 44% compared to 40% in 2010 and 2011 combined). Reporting of data from health facilities can be heavily influenced with lack of registers and lack of training to staff responsible for compiling and preparing aggregate summaries. Timeliness of reporting can also be affected by lack of regular transportation from facilities to district offices. Overall, HMIS data reporting varies across study districts as it also depends on the number and type of facilities in the district.

Contribution of health information system on the strength of FANC and EmOC programme implementation relies on the use of information generated from health facilities and from population-based surveys. Information use indicates how important district planners and service providers base their decision-making on information. While the study obtained sufficient data on the HMIS reporting rate and timeliness and on some information regarding districts' ICT capacity, it was difficult to obtain documents with evidence on information use for planning of maternal health programmes. It was not possible in most districts to access the district's main official document referred to as the comprehensive council health plan (CCHP). CCHP document presents a detailed record of district information on health expenditure, which is considered sensitive and rarely given for research purposes. Important to this study, the document also contains the district's performance indicators for a following year's planned implementation of health programmes. Using only the number of reports submitted to the district medical officer's office and

the timeliness of reporting only captures part of the picture. An assessment of the health information system would have been complete with an examination of how district planners incorporate the collected information to prioritize maternal health programmes in their respective districts.

Leadership and governance: This study found that the majority of the district health management teams conduct quarterly visits, with some districts conducting monthly visits. A follow up investigation of why districts had different frequencies in conducting supportive supervision visits found two main reasons: districts conducting monthly visits were more likely to do so due to set priorities and that districts depended on sufficiency in funds and availability of vehicles. The World Health Organization defines supervision as "the support and guidance that a supervisor gives staff for whom he or she is responsible in order for them to perform their duties effectively, competently, and receive job satisfaction".[172] In conducting regular supportive supervision visits to health facilities in respective districts, district health management teams improve the performance of health workers and the overall performance of health facilities. Supportive supervision include district supervisors assisting staff in charge of facilities and clinicians in setting goals, putting in place the mechanisms for achieving annual targets, and monitoring and evaluating overall facility performance. In completing the supervision task, visiting officials provide feedback to the management for effective planning and execution of activities.

4.5 Conclusion

While government efforts can be appreciated in producing a reasonable number of nurses and midwives, more efforts are still required in increasing the number of doctors and other non-physician clinicians. Efforts to increase production of the health workforce from colleges and universities should be accompanied by an equitable distribution of the human resources to mitigate observed differences between rural and urban districts. In addition, in order to improve FANC and EmOC service delivery, district health authorities and the Tanzanian government in general need to build more health facilities that are equitably distributed in rural and urban areas, facilities that are equipped (with operating theatres and blood banks) as well as facilities that are staffed with skilled personnel. Availability of essential FANC and EmOC drugs and supplies should be given priority by both the

central and local governments adequately addressing the issues known to affect continued stocking. Essential drugs and supplies are important in ensuring that women in need of the services receive them and help in avoiding unnecessary inconveniences which, for some, this could result in fatal effects.

The Tanzanian government should also find more ways to increase funding to districts and in expanding financial protection to the most vulnerable including enforcing the fee-exemption to maternal health services as it should be. Further, while the Tanzanian Ministry of Health and Social Welfare is scaling up the DHIS system across all districts in Mainland Tanzania, much work is needed to ensure the DHIS system functions optimally at the district level. Improving the DHIS operability should be accompanied with improving health information system infrastructures in lower facilities. Full functioning of the HMIS from lower facilities upwards will ascertain feeding of information to districts and upper levels of the health system thereby facilitating continued production of the evidence required in planning FANC and EmOC health services. While council health management teams prepare their annual plans and budgets to support supervision visits to health facilities in districts, there is a need to streamline their routes and time spent in health facilities in order to maximise their efforts providing the planned support to low-level facilities for gradual improvement in the provision of FANC and EmOC services.

Scaling the implementation of maternal health programmes in Tanzania (and in low- and middle-income countries alike) requires that all six building blocks operate to their full capacity in contributing to the overall strength of FANC and EmOC programmes. Inequity in the implementation strength between urban and rural districts still remains a visible divisive factor. Meanwhile, an agenda remains for further work on how advocacy groups in districts can use these results to hold authorities accountable in improving maternal health services and help reduce the effects due to poor or unavailable services.

Chapter 5: Weighted Scores of the Implementation Strength

5.1 Introduction

Composite indices have been used in diverse fields as tools for comparison, policy analysis, and public communication.[105] [173] For example, the Dow Jones Industrial Average and the Consumer Price Index are composite indices that have become highly important measures for discussions of, respectively, investment forecasting, and markets of consumer goods and services.[174] [175] In the health sector, several composite indices, such as the World Health Organization's Global Burden of Disease disability weights and the Health System Performance Index, have been used to benchmark countries on numerous indicators of health.[176] [177] In general terms, composite indices share three common features: they are usually multidimensional in nature, they can measure specific variables of interest at national level, and they are designed to enable comparison across different units.[178] Composite indices can also be applied at the global level or at national or subnational level to stimulate a "programmatic response from lower level administrative units".[178] Advantages and disadvantages of composite indicators over other indicators are presented later in the Chapter under discussion.

This chapter begins by highlighting the steps used to construct composite scores, followed by a description of the methods used to aggregate constituent scores and on how the weighting scheme was applied on the unweighted summary scores. The chapter also discusses how uncertainty around overall composite scores was estimated and how reliability of the index was assessed. A summary of unweighted scores from Chapter 4 (Tables 4.1 to 4.7) and those calculated for this Chapter (the overall weighted scores) are presented and disaggregated by components and by study districts. Maps of study districts are used to show variation of implementation strength scores in three district categories: low performing districts, medium performing districts and high performing districts. The chapter concludes by highlighting some methodological challenges. It is important to note that implementation strength scores presented in this chapter are intended to assess the

inputs and processes as they relate to the "dose" side of the conceptual framework presented earlier in Chapter 2 on Methods Overview Section 2.6.

5.2 Summary steps for constructing a composite index

It was highlighted earlier (Section 1.4) that, construction of composite indices can be a complicated process and that the process incorporates several methodological concerns which would require care in addressing them. It is therefore necessary to be thoughtful of the likely challenges that come along with computing composite indices and in preparing the rankings. In search of a valid procedure for use in preparing the FANC and EmOC programmes composite scores and the rankings of study districts, two guidelines were examined for use in this study: the guidelines prepared by the OECD, "Handbook on Constructing Composite Indicators"[105] and the guidelines outlined by Jacobs et al's paper, "Measuring performance: An examination of composite performance indicators".[100] Table 5.1 presents the two sets of guidelines in form of steps. The two guidelines seemed similar to a large extent except in few steps. Due to their close similarities, this study adopted using the OECD guidelines mainly because of its conceptual appeal. Elsewhere, Smith et al had pointed other steps for constructing composite indicators for the health system, which were also largely similar to those by Jacobs et al and included: selection of the indicators, a consideration/assessment of the collinearity of the components, identification of weights for the components, transforming the constituent indicators, environmental influences on performance, and the analytic approaches to inferring efficiency.[100]

Table 5.1: Steps/guidelines for building a composite indicator

Steps by the OECD handbook	Steps by Jacobs et al
1. Theoretical framework Provides the basis for the selection and combination of variables into a meaningful composite indicator under a fitness-for-purpose principle (involvement of experts and stakeholders is envisaged at this step).	1. Choosing the units to be assessed The choice of the organisational units (such as primary or secondary care) will have an impact on the choice of indicators which are available with which to assess their performance
2. Data selection Should be based on the analytical soundness, measurability, country coverage, and relevance of the indicators to the phenomenon being measured and relationship to each other. The use of proxy variables should be considered when data are scarce (involvement of experts and stakeholders is envisaged at this step).	2. Choosing the organisational objectives to be encompassed in the composite A decision must be made about the overall objectives against which the organisations are to be assessed (This is probably mainly a political task)
3. Imputation of missing data Is needed in order to provide a complete dataset (*e.g.* by means of single or multiple imputation).	3. Choosing the indicators to be included Most fundamentally important where judgement is required. A different set of indicators will produce a different composite indicator and hence a different set of rankings, although it is not known how different.
4. Multivariate analysis Should be used to study the overall structure of the dataset, assess its suitability, and guide subsequent methodological choices (*e.g.*, weighting, aggregation).	4. Transforming measured performance on individual indicators There is no need for any transformation if it is possible to specify a weight that indicates the relative value to the composite of an extra unit of attainment in that dimension *at all levels of attainment*. Otherwise a transformation is required.
5. Normalisation Should be carried out to render the variables comparable.	5. Combining the indicators The different dimensions of performance measured on different scales (which are then transformed into a common scale) then need to be combined in a meaningful way.
6. Weighting and aggregation Should be done along the lines of the underlying theoretical framework.	6. Adjusting for environmental or uncontrollable factors Some units must operate in more adverse environmental circumstances which may make the attainment of performance outcomes more difficult for them. Thus for a given level of expenditure, the production possibility frontiers for these systems will lie inside those with more favourable environmental conditions
7. Uncertainty and sensitivity analysis Should be undertaken to assess the robustness of the composite indicator in terms of e.g., the mechanism for including or excluding an indicator, the normalisation scheme, the imputation of missing data, the choice of weights, and the aggregation method.	7. Examining variations in efficiency Regulators may be interested in exploring the efficiency with which organisations use resources in relation to achieving the performance measured on the composite. This leads to the examination of performance in relation to some measure of resource use, usually cost.
8. Back to the data Is needed to reveal the main drivers for an overall good or bad performance. Transparency is primordial to good analysis and policymaking.	8. Sensitivity analysis of the construction of the composite indicator The outcomes and rankings of individual units on the composite may largely depend on the decisions taken at each of the preceding steps. As such, an important consideration is the use of a sensitivity analysis to explore the robustness of rankings to the inclusion and exclusion of certain variables, changes in the weighting system, using different transformation methods and setting different decision rules to construct the composite
9. Links to other indicators Should be made to correlate the composite indicator (or its dimensions) with existing (simple or composite) indicators as well as to identify linkages through regressions.	9. Empirical analysis ***First part:*** exploratory – examines the underlying nature of the data, including the distributions of the underlying indicators, how they have been transformed, the correlations between the indicators and a factor analysis. ***Second part***: use of the dataset to construct a new composite index through a simulation exercise.
10. Visualisation of the results Should receive proper attention, given that the visualisation can influence (or help to enhance) interpretability	

5.3 Methods

5.3.1 Step 1: Theoretical framework

The basis for selecting the indicators of the FANC and EmOC programme components was pointed out in Chapter 3. Indicator selection was preceded by a conceptual framework whose central objective was to delineate the interrelationship between programme strength and programme results and the contribution of contextual factors in either programme implementation or programme results. It was argued in Chapter 3 that, programmatically, the WHO health-system-building blocks appeared to package well the essential functional parts of a large-scale maternal health programme such as the FANC and EmOC programmes. Choice of such components is most likely to influence the choice of indicators.[190] To minimise the chances for selecting unrelated components, a literature search was conducted (Chapter 3) resulting into selection of the building blocks. Chapter 3 further described the multidimensional nature of the programme components (blocks) – a basis on which each programme component was examined to identify its constituent activities and/or elements. To validate the selection of activities and elements, maternal health experts reviewed them and provided a final list. Based on the final list of components and activities/elements, indicators for each programme component were selected. To estimate the implementation strength of the two programmes, a composite indicator is now required that will enable combining each programme component's contribution to implementation strength.

5.3.2 Step 2: Selection of indicators

Step one (of the OECD handbook) on theoretical framework concluded with a selection of proxy indicators that were used for collecting programme data. Indicators for the implementation strength were first introduced in section 2.3 and briefly discussed in section 2.6.5 (both sections of Chapter 2), and extensively discussed in Chapter 4 (section 4.2.2). A summary of the indicators for the implementation strength is shown in Table 2.3 and included the following: for human resources: the number of health workers per 10,000 population); for drugs and supplies: availability of tracer drugs; service delivery: number of health facilities per 10,000 population, the number of inpatient beds per 10,000 population, and number of first antenatal care visits per 10,000 population; for health information

system: quarterly HMIS reporting rate and timeliness of HMIS data; for health financing: per capita district revenues and expenditure; and for leadership and governance: number of supportive supervision visits to health facilities. As pointed out in previous sections, indicators for the implementation strength were included into the composite index based on their relevance in measuring components, their use popularity in the literature, suggestions by the World Health Organisation, and most importantly, availability of their data in study districts. Choice of the indicators culminated in exclusion of other potential indicators for the implementation strength. While this can have effect in the resulting composite index, Jacobs et al concluded that composite indicators "can only draw on data that is available".[100] It is also possible that a composite indicator can only be fit-for-purpose depending on the data quality of its constituent indicators, and other issues such as those pointed out by Jacobs et al as levels of aggregation, choice of different types of indicator and the incentive effects of the composite indicators. Data quality and reliability issues for the selected indicators of implementation strength were discussed in Chapter 4 (Section 4.3).

5.3.3 Step 3: Imputing missing data

Data for most indicators of the implementation strength were available from identified sources. However, some districts had missing data for the components of health financing and health information system. For example, on district revenues, some districts had missing data on development funds, MMAM funds, Global Funds, cost sharing, district own funds and insurance funds (Table 4.4). For the health information system component, three districts did not have information on either reporting or timeliness of HMIS reports (Table 4.6). Due to extreme differences in revenues between study districts (for example, on Global Fund, Kilosa district had TZS=4,389,316,500 while Iringa Urban had TZS=5,054,800), all missing data on district revenues for urban districts were imputed with the median value for urban districts and rural districts with the median value for rural districts. All three rural districts with missing data on health information system were imputed with values of the arithmetic mean for rural districts.

5.3.4 Step 4: Multivariable analysis using PCA

The main aim of performing multivariable analysis is to predict values of one explanatory variable given the value of one or more explanatory variables in a regression model. Following on the OECD handbook, in order to explore the data and to check for the underlying structure of the relationship between the composite scores of programme components, the study used Principal Component Analysis method (PCA) due to its ability to generate leading principal components. PCA enhances clustering of structure from which features with the most variance are the most relevant to the clustering.[179] [180] PCA is a mathematical method that creates linear combinations of the variables that are independent of each other. The first principal component is the linear combination of the variables that gives the greatest variability, and the last component reflects the least. The scores for each component reflect correlations between the variables. Variables that are closely correlated with each other will have high scores on at least one principal component. Even though all the data (weighted scores) were expressed in the same unit, percentages, PCA analysis was run on the correlation matrix in order to ensure equal weight.[181] The numbers of principal components extracted were those whose associated eigenvalues were greater than one.[182]

5.3.5 Step 5: Normalising data

Due to differences in units of measurements for some of the indicators of a particular programme component, normalisation of data was conducted in order to make constituent indicators comparable. There are a number of ways of normalising data including the Min-Max and Distance to a reference.[100] [183] To normalise data for the indicators of the implementation strength, Chapter 4 (Section 4.2.4) discussed how the Min-Max method was used. Briefly, the Min-Max method normalises data in indicators to have an identical range of values from 0% to 100% by using the formula $(X_i - X_{min}) / (X_{max} - X_{min})$ where X_i is the value for the ith district (i = 1, 2, 3, …, 23) and X_{min} and X_{max} are the minimum and maximum values of X_i respectively.

5.3.6 Step 6: Aggregating scores and applying the weights

To aggregate constituent scores of a programme component in each district, the summary scores from the unweighted indicators were calculated (Chapter 4 on unweighted scores). The summary scores were obtained by calculating the arithmetic means of constituent indicators. Where arithmetic means were not on the same scale, they were normalised to measure on the same scale of percentages. For example, the summary score for the EmOC' drugs and supplies component was calculated by adding up the percentage availability in surveyed health facilities of the following indicators (all in the same unit, %): delivery packs, episiotomy scissors, manual vacuum extractors, vacuum aspiration or dilation and curettage kits, neonatal bags and masks, injectable antibiotics, injectable uterotonic, and injectable magnesium sulphate – and dividing by the number of the indicators, eight (all variables/indicators had complete dataset). Another example for this was in obtaining the summary score for the health financing component whose indicators on revenues and on expenditures were separately added and divided by the number of district population to obtain their corresponding district's per capita revenues and per capita expenditure. Per capita results were further converted to district scores using the Min-Max formula to obtain scores in percentage and averaging them to obtain an overall district score for the health financing component. District revenues included indicators on Cost Sharing, Insurance Funds, Basket Funds, Block Grants, Development Funds, District Own Funds, MMAM Funds, Global Fund whereas district expenditure included recurrent expenditure and expenditure on development activities. The arithmetic mean is an appropriate aggregation method for indicators measured on same scale. The alternative aggregation method would be using geometric means.

Having generated the summary scores for all component indicators, the next step was to adjust components according to assigned weights. This was done by applying the weighting scales obtained from maternal health experts, which were (for FANC and EmOC respectively): human resources (23.8% and 24.8%), drugs and supplies (19.7% and 19.8%), service delivery (17.2% and 18.6%), health financing (17.5% and 16.4%), health information system (11.2% and 10.3%) and leadership and governance (10.6% and 10.1%). The overall implementation strength measure

(overall composite score) was defined as a linear aggregate of the six components such that:

$CompScore_{FANC}$ = 23.8%HR + 19.7%$D\&S_{FANC}$ + 17.2%SD_{FANC} + 17.5%HF + 11.2%HIS + 10.6%L&G

$CompScore_{EmOC}$ = 24.8%HR + 19.8%$D\&S_{EmOC}$ + 18.6%SD_{EmOC} + 16.4%HF + 10.3%HIS + 10.1%L&G

Where:

$CompScore_{FANC}$=Overall Implementation Strength Score for FANC
$CompScore_{EmOC}$=Overall Implementation Strength Score for EmOC
HR=Human resources
$D\&S_{FANC}$=Drugs and Supplies for FANC
$D\&S_{EmOC}$=Drugs and Supplies for EmOC
SD_{FANC}=Service Delivery for FANC
SD_{EmOC}=Service Delivery for EmOC
HF=Health Financing
HIS=Health Information System
L&G=Leadership & Governance

[note: both FANC and EmOC programmes shared the same scores on HR, HF, HIS, and L&G components except for the D&S and SD for which each had own component scores]

The Stata command sum was used to generate the overall district's composite score (mean) in conjunction with the option iweight (importance weight) that adjusted the district's component scores according to their assigned weights. For example, the overall composite score for Geita district for FANC programme was calculated as:

sum $Geita_scores_{FANC}$ [iweight=FANCweights]

Where:
$Geita_scores_{FANC}$=Unweighted summary scores for each FANC component

FANCweights=Respective weights assigned to FANC components.

The same approach was applied for the EmOC programme overall composite score, using the respective EmOC scores and weights. Overall, absolute composite scores were grouped to categorise districts into three performance levels: low performing districts (with scores less than the lower quartile boundary), medium performing districts (with scores between the lower and the upper quartile boundaries), and high performing districts (those with scores higher than the upper quartile boundary).

5.3.7 Step 7: Uncertainty and Sensitivity Analysis

In developing the composite indices, possible uncertainties could have been introduced because of carrying out some processes based on subjective judgement and sample data. These could have been due to: selection of indicators, choice of the weighting approach (using maternal health experts), the quality of programme data, aggregation of scores from individual indicators, rescaling scores (normalisation) and so forth. Performing sensitivity analysis and estimating uncertainty can "help gauge the robustness of the composite indicator and improve transparency".[105] The robustness of the composite indicator determines whether findings obtained by changing variable levels are consistent with original findings, and that the conclusions are similar.[184] To analyse the robustness of the composite index, first scatter plots were used to assess the relationship between each component and the overall composite score, followed by successfully adding one component after the other on the FANC and EmOC regression models discussed in Sections 5.3.6 and 5.3.7. Statistical significance was assessed based on the regression coefficients with corresponding 95% confidence intervals. Using the final models, sensitivity analysis was conducted to assess the effect programme components had on the overall composite score of either programme (FANC or EmOC). A sensitivity parameter γ was defined as the effect on overall implementation strength score resulting from a δ percentage change in one component while controlling for other components at their modelled values. Four values of δ: 25%, 50%, 75% and 100% (0% = base-case) were used to determine sensitivity changes γ in the overall implementation strength score. Results were presented in bar graphs showing sensitivity indexes γ with associated δ percentage changes.

In addition, the reliability of the composite index was assessed by using Cronbach's alpha scores.[185] [186] Cronbach's alpha scores test the 'internal consistency' of a scale, a quantitative measure of the degree to which items are related to each other. According to Strainer and Norman, a scale is said to be homogenous if "all of the items are tapping different aspects of the same attribute being measured and not different parts of different traits".[187] Ponterotto and Ruckdeschel[188] proposed a value of alpha=0.70 to be 'good' for a scale with fewer than seven items and evaluated with fewer than 100 subjects – which was relevant for this study. Cronbach's alpha scores were presented in a table showing both inter-component correlations and the overall internal consistency score.

5.3.8 Step 8: Back to the data (Ranking districts)

To identify programme components responsible for an overall good or bad performance in study districts, study districts were ranked based on their composite scores for each of the two programmes. Two tables (one for each programme) were prepared showing study districts with their corresponding unweighted and weighted scores. Maps were generated using ArcGIS 10.2 for Desktop (ESRI, United States) to display study districts by low-, medium-, and high- performing groups. Low-performing districts were those with a composite score of up to the lower quartile score; medium-performing districts were those with scores between the lower quartile and the upper quartile; and high-performing districts were those with scores above the upper quartile.

5.3.9 Step 9: Link to other indicators

Step 9 of the guidelines was separately completed in Chapter 8 on dose-response by linking the composite indices with other indicators through regression analyses. Regressions used data from programme strength, coverage and contextual factors.

5.3.10 Step 10: Visualisation of the results

To visualise the results, tables, graphs and maps have been produced to help interpretability of the results. The following sections present the results of the composite indicators disaggregated by study districts.

5.4 Results

5.4.1 Overall scores and district rankings

Tables 5.2 and 5.3 respectively show the FANC and EmOC implementation strength aggregate scores of the programme components, with corresponding ranks of study districts. The tables also show the overall composite scores for both unweighted and weighted components and the corresponding weighted overall ranks of study districts. Table 5.2 shows that Arusha Urban was the strongest district in implementing FANC programme activities, with an overall weighted composite score of 86%. Musoma Rural district had the least FANC overall weighted score of 24%.

On individual programme components for FANC, Arusha Urban had the highest composite scores on four of the six programme components. The district scored 100% on human resources component, 96% on drugs and supplies, 94% on health financing, and 77% on service delivery. The remaining programme components were topped by the rural districts of Geita with 55% on the health information system component and Bagamoyo with 100% on the leadership and governance component. Table 5.3 shows Arusha Urban being the district with the strongest implementation of the EmOC programme, with an overall weighted composite score of 80%. Musoma Rural was the weakest district in implementing the EmOC programme, with an overall weighted composite score of 22%. In addition, in implementing the EmOC programme, Arusha Urban was the leading district in the three components of human resources, drugs and supplies, and health financing. The other three components were topped by Iringa Urban district with 71% on service delivery, Geita with 55% on health information system and Bagamoyo with 100% on leadership and governance.

Table 5.2: Overall Composite Scores and Ranking of Study Districts on FANC Programme Implementation Strength

	Study District	Human resources		Service Delivery		Drugs & Supplies		Health Financing		Health Information System		Leadership and Governance		Overall FANC Score		
		Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Unweighted	Weighted	Rank
Urban Districts	Arusha Urban	100%	1	77%	1	96%	1	94%	1	46%	3	76%	11	82%	86%	1
	Ilala	38%	4	75%	2	69%	6	59%	4	28%	14	68%	14	56%	56%	3
	Iringa Urban	30%	5	67%	3	90%	4	66%	3	41%	6	79%	6	62%	61%	2
	Kinondoni	30%	7	49%	11	65%	9	40%	10	19%	23	54%	20	43%	43%	10
	Mtwara Urban	14%	13	59%	4	80%	5	57%	5	35%	11	88%	4	55%	52%	5
	Songea Urban	66%	2	57%	5	36%	16	27%	15	23%	18	88%	5	49%	49%	7
	Tanga Urban	17%	12	51%	8	67%	7	70%	2	24%	17	52%	21	47%	46%	8
	Temeke	17%	11	38%	14	92%	3	54%	6	45%	4	75%	12	54%	51%	6
Rural Districts	Babati	18%	10	51%	9	30%	20	27%	14	36%	10	96%	3	43%	38%	13
	Bagamoyo	42%	3	54%	7	64%	10	38%	12	39%	7	100%	1	56%	54%	4
	Geita	13%	15	38%	13	50%	12	12%	20	55%	1	97%	2	44%	38%	12
	Kahama	9%	18	29%	20	60%	11	10%	22	48%	2	48%	22	34%	31%	16
	Kasulu	7%	20	50%	10	32%	19	14%	18	22%	19	47%	23	29%	26%	22
	Kilosa	29%	8	33%	17	12%	23	44%	7	27%	15	67%	15	35%	33%	15
	Kondoa	2%	22	31%	18	39%	15	23%	16	38%	8	63%	16	33%	29%	19
	Mbozi	14%	14	39%	12	18%	22	14%	19	28%	12	79%	9	32%	28%	20
	Moshi Rural	30%	6	37%	15	66%	8	42%	8	22%	20	75%	13	45%	44%	9
	Muleba	7%	21	17%	22	35%	18	28%	13	37%	9	79%	7	34%	29%	18
	Musoma Rural	8%	19	16%	23	43%	14	3%	23	19%	22	79%	8	28%	24%	23
	Ruangwa	22%	9	31%	19	35%	17	39%	11	42%	5	79%	10	41%	37%	14
	Singida Rural	11%	16	37%	16	92%	2	22%	17	28%	13	62%	18	42%	41%	11
	Sumbawanga R	9%	17	55%	6	22%	21	11%	21	19%	21	63%	17	30%	27%	21
	Uyui	0%	23	28%	21	46%	13	42%	9	27%	16	55%	19	33%	30%	17
	Overall Mean Score	23%		44%		54%		36%		32%		73%		44%	41%	
	Standard Deviation	22%		17%		25%		23%		10%		16%		13%	14%	

Table 5.3: Overall Composite Scores and Ranking of Study Districts on EmOC Programme Implementation Strength

	Study District	Human resources		Service Delivery		Drugs & Supplies		Health Financing		Health Info System		Leadership & Gov		Overall Composite Score		
		Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Unweighted	Weighted	Rank
Urban Districts	Arusha Urban	100%	1	69%	2	71%	1	94%	1	46%	3	76%	11	76%	80%	1
	Ilala	38%	4	65%	4	62%	5	59%	4	28%	14	68%	14	53%	53%	4
	Iringa Urban	30%	5	71%	1	68%	2	66%	3	41%	6	79%	6	59%	57%	2
	Kinondoni	30%	7	45%	11	46%	15	40%	10	19%	23	54%	20	39%	39%	13
	Mtwara Urban	14%	13	68%	3	55%	9	57%	5	35%	11	88%	4	53%	49%	5
	Songea Urban	66%	2	65%	5	60%	7	27%	15	23%	18	88%	5	55%	56%	3
	Tanga Urban	17%	12	63%	6	61%	6	70%	2	24%	17	52%	21	48%	47%	7
	Temeke	17%	11	32%	14	55%	11	54%	6	45%	4	75%	12	46%	42%	9
Rural Districts	Babati	18%	10	52%	8	41%	18	27%	14	36%	10	96%	3	45%	40%	11
	Bagamoyo	42%	3	50%	9	39%	20	38%	12	39%	7	100%	1	51%	48%	6
	Geita	13%	15	7%	23	41%	19	12%	20	55%	1	97%	2	38%	30%	17
	Kahama	9%	18	10%	20	50%	13	10%	22	48%	2	48%	22	29%	25%	20
	Kasulu	7%	20	35%	12	36%	21	14%	18	22%	19	47%	23	27%	25%	22
	Kilosa	29%	8	28%	17	50%	12	44%	7	27%	15	67%	15	41%	39%	12
	Kondoa	2%	22	27%	18	55%	10	23%	16	38%	8	63%	16	35%	31%	16
	Mbozi	14%	14	28%	16	30%	23	14%	19	28%	12	79%	9	32%	28%	19
	Moshi Rural	30%	6	53%	7	64%	3	42%	8	22%	20	75%	13	48%	47%	8
	Muleba	7%	21	25%	19	47%	14	28%	13	37%	9	79%	7	37%	32%	15
	Musoma Rural	8%	19	8%	22	42%	17	3%	23	19%	22	79%	8	27%	22%	23
	Ruangwa	22%	9	33%	13	58%	8	39%	11	42%	5	79%	10	46%	42%	10
	Singida Rural	11%	16	30%	15	62%	4	22%	17	28%	13	62%	18	36%	33%	14
	Sumbawanga R	9%	17	47%	10	36%	22	11%	21	19%	21	63%	17	31%	28%	18
	Uyui	0%	23	9%	21	43%	16	42%	9	27%	16	55%	19	29%	25%	21
	Overall Mean Score	23%		40%		51%		36%		32%		73%		43%	40%	
	Standard Deviation	22%		21%		11%		23%		10%		16%		12%	14%	

Figures 5.1 through 5.8 are maps of Tanzania showing study districts in the three strength categories: poor performing districts, medium performing districts and high performing districts, based on FANC and EmOC composite scores for each programme component. Except for the components of service delivery (Figures 5.3 and 5.4, FANC) and drugs and supplies (Figures 5.5 and 5.6, EmOC), all other maps show same scores for FANC and EmOC programmes. Figure 5.1 shows study districts in strength categories for the human resources component, with Arusha Urban, Moshi Rural, Bagamoyo, Ilala, Iringa Urban, and Songea Urban being in the high performing group of districts. Study districts scored between 0% and 100% for the component. Districts in the low performing group scored between 0% and 9% and consisted of six districts of Musoma Rural, Muleba, Kasulu, Kahama, Uyui, and Kondoa. Figure 5.2 shows study district performance on the health financing component. Six four districts were in the high performing group, all of which were urban districts of Arusha Urban, Tanga Urban, Iringa Urban, Ilala, Mtwara Urban, and Temeke. Musoma Rural was the lowest scoring district with a score of 3% while Arusha Urban had the highest score of 94% for the health financing component.

Figure 5.3 and Figure 5.4 show the overall performance of study districts on the service delivery component for FANC and EmOC programmes respectively. High performing districts on service delivery for FANC programme scored between 55% and 77% and included Arusha Urban, Ilala, Iringa Urban, Mtwara Urban, Songea Urban, and Sumbawanga Rural. High performing districts for EmOC service delivery scored between 52% - 72% and included the districts Iringa Urban, Arusha Urban, Mtwara Urban, Ilala, and Songea Urban. The low performing group of study districts on FANC service delivery included the districts of Muleba, Uyui, Ruangwa, Kondoa, Kahama, and Musoma Rural while those in the low performing group on EmOC service delivery included Geita, Kahama, Muleba, Musoma Rural, and Uyui districts. Figure 5.5 and Figure 5.6 show study districts on drugs and supplies for FANC and EmOC programmes respectively. The high performing group of study districts in stocking drugs and supplies for the FANC programme had six districts – four of which were also in the high performing group for stocking EmOC drugs and supplies. The high performing group on FANC drugs and supplies included one rural district of Singida Rural and five urban districts of Temeke, Ilala, Kinondoni, Arusha Urban, Iringa Urban, and Mtwara Urban, all of which scored between 69% and 96%. High performing group of study districts on EmOC drugs and supplies scored between 62% and 71% and included two rural districts of Singida Rural and

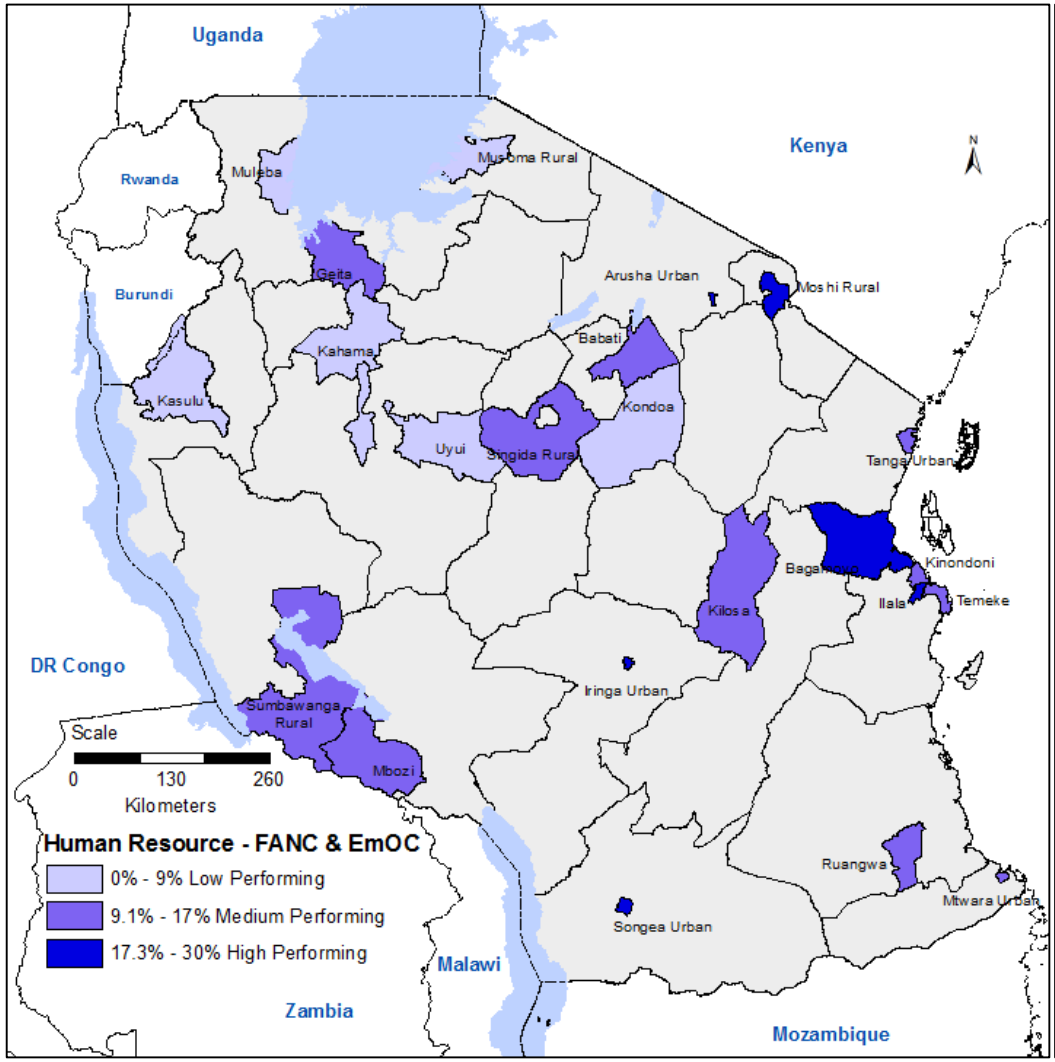


Figure 5.1: Study districts performance on FANC & EmOC HR

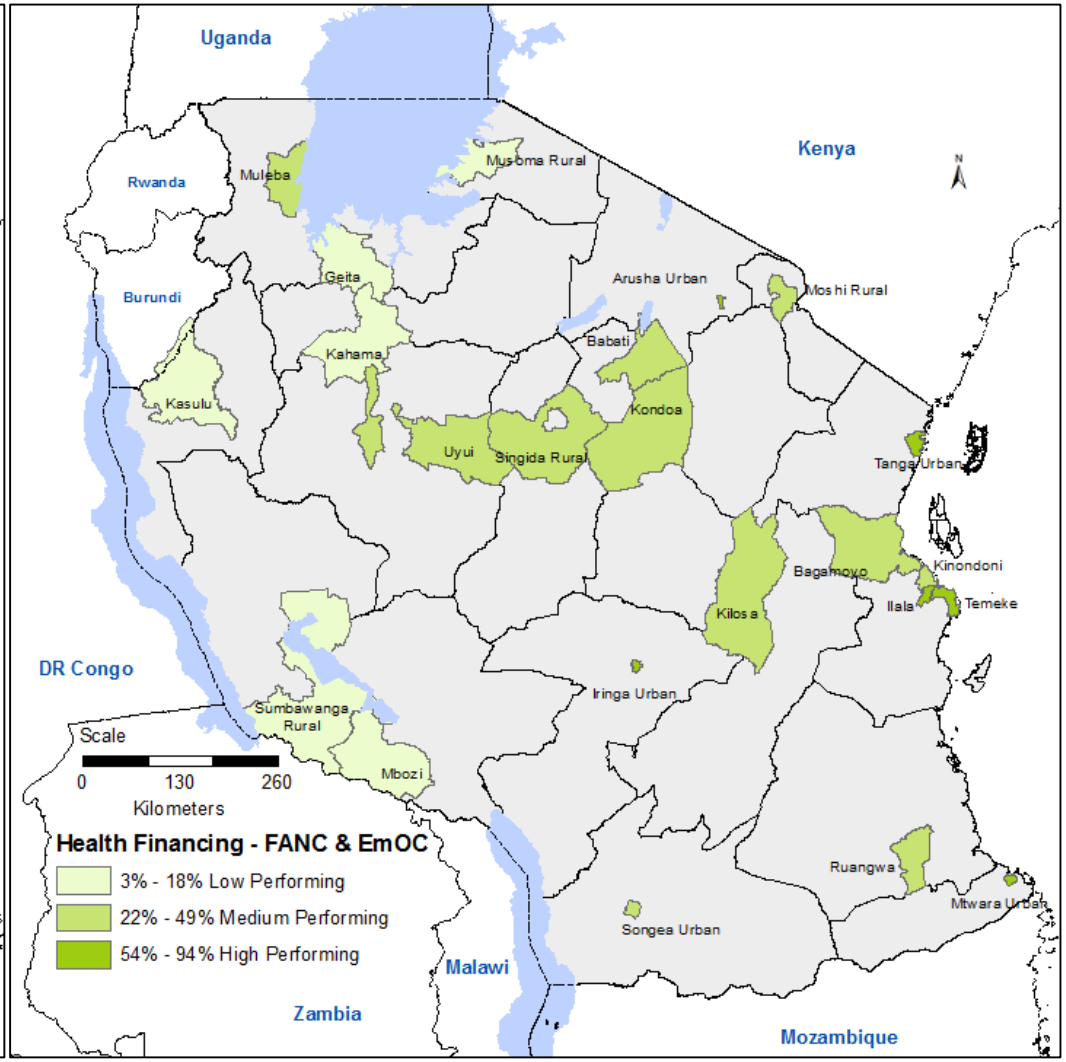


Figure 5.2: Study districts performance on FANC & EmOC Financing

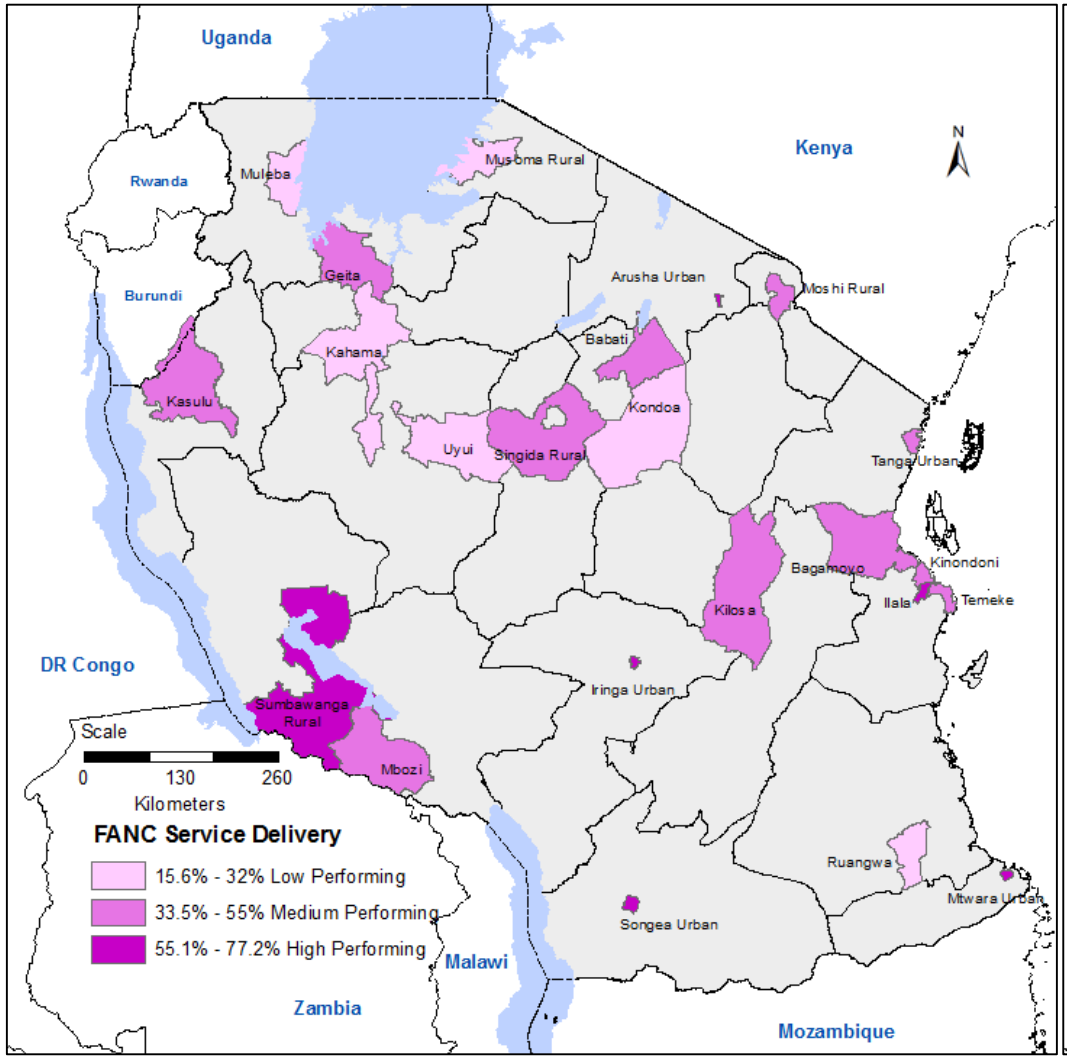


Figure 5.3: Study districts performance on FANC Service Delivery

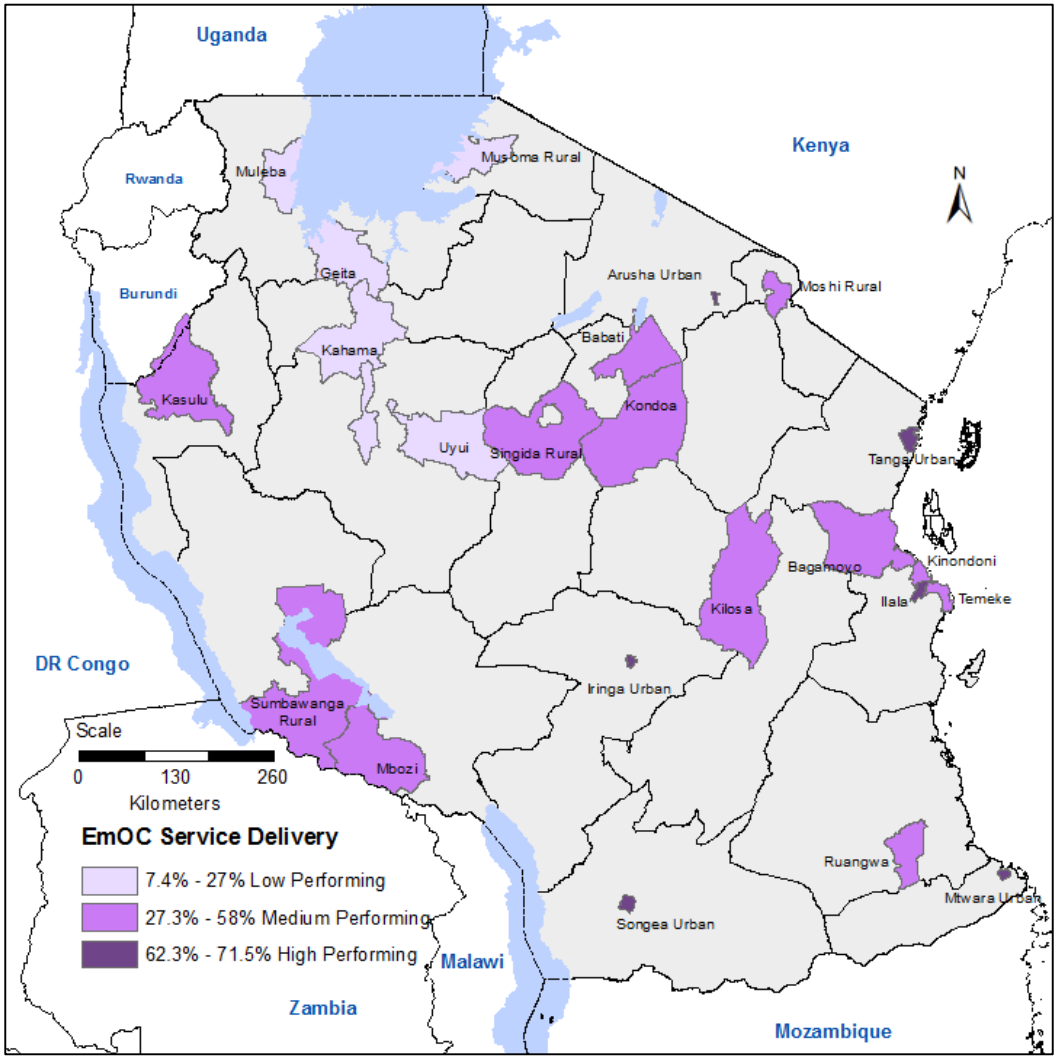


Figure 5.4: Study districts performance on EmOC Service Delivery

Figure 5.5: Study districts performance on FANC Drugs & Supplies

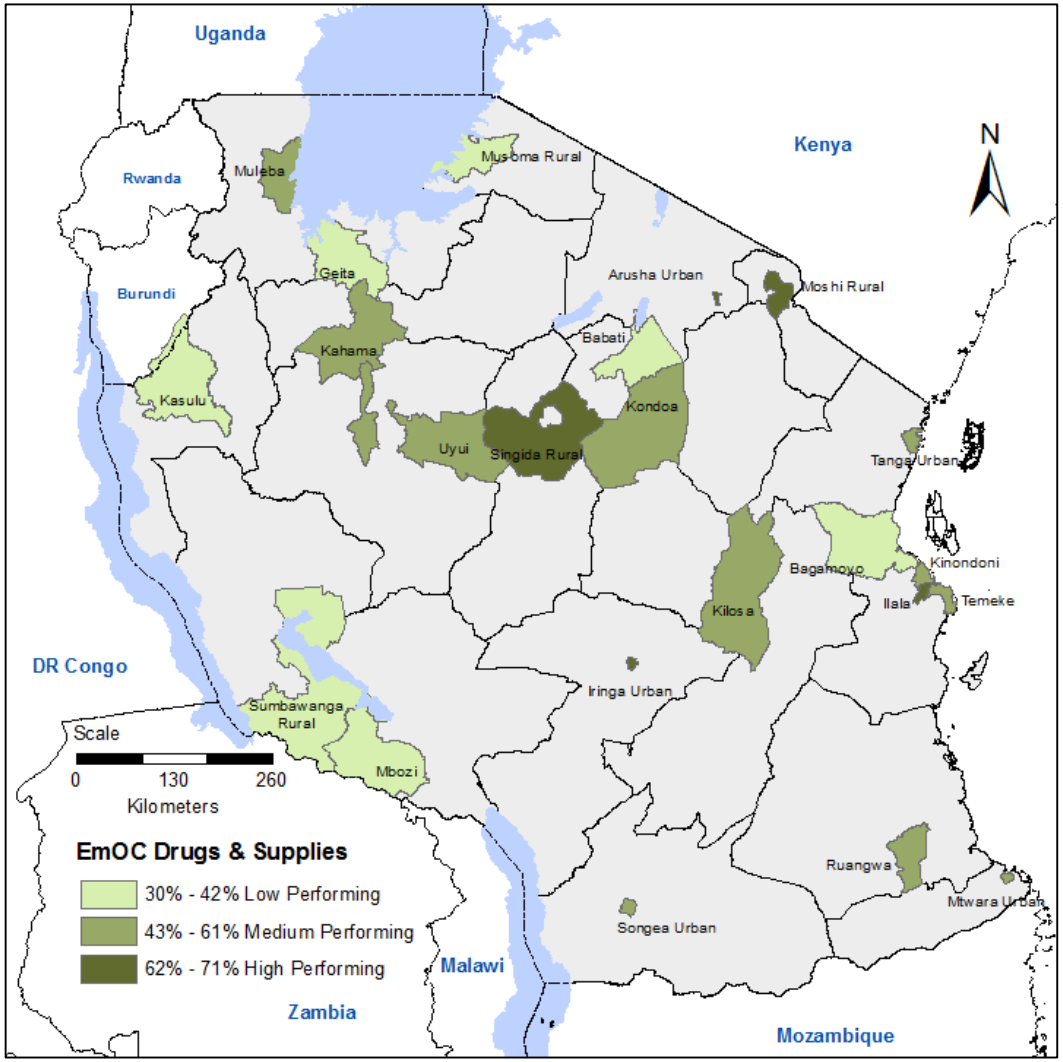


Figure 5.6: Study districts performance on EmOC Drugs & Supplies

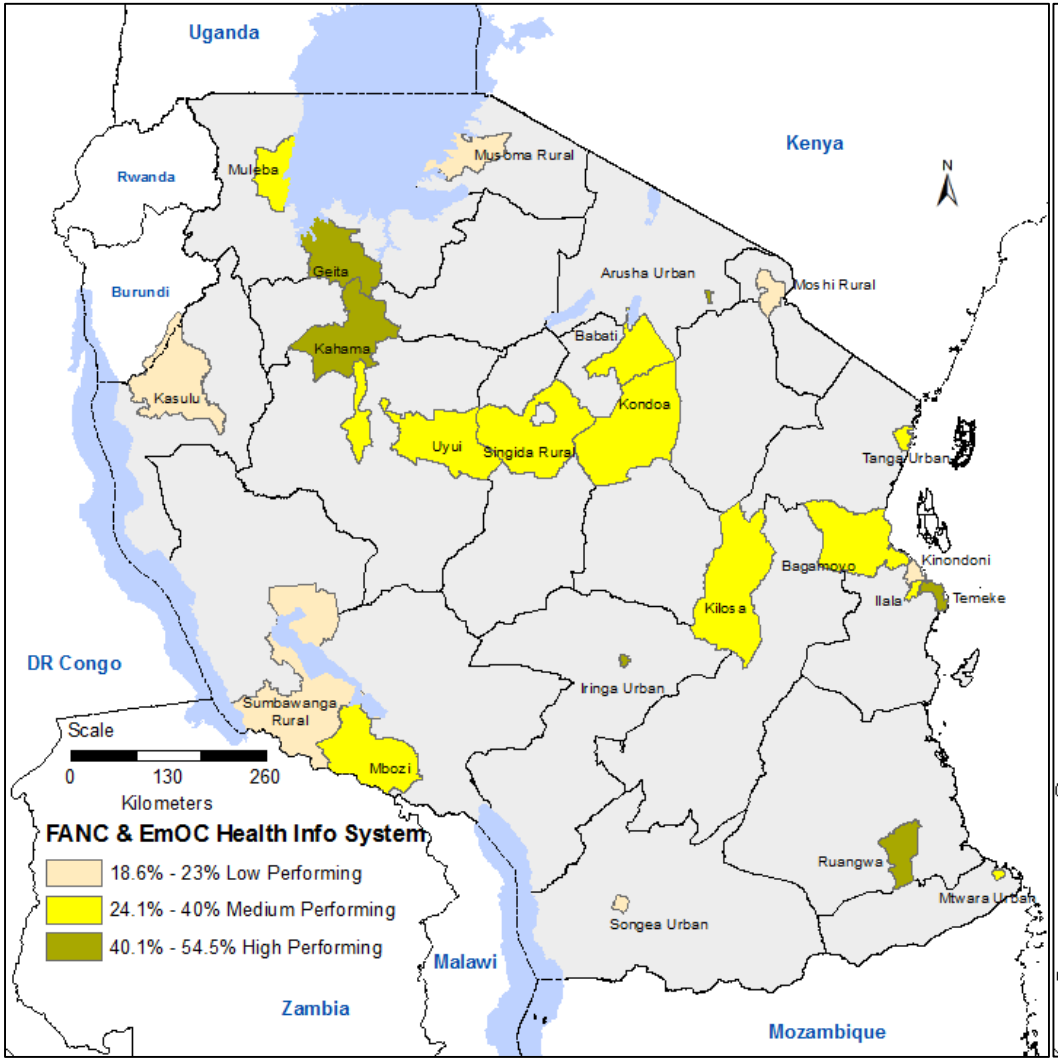


Figure 5.7: Performance on FANC & EmOC Health Info System

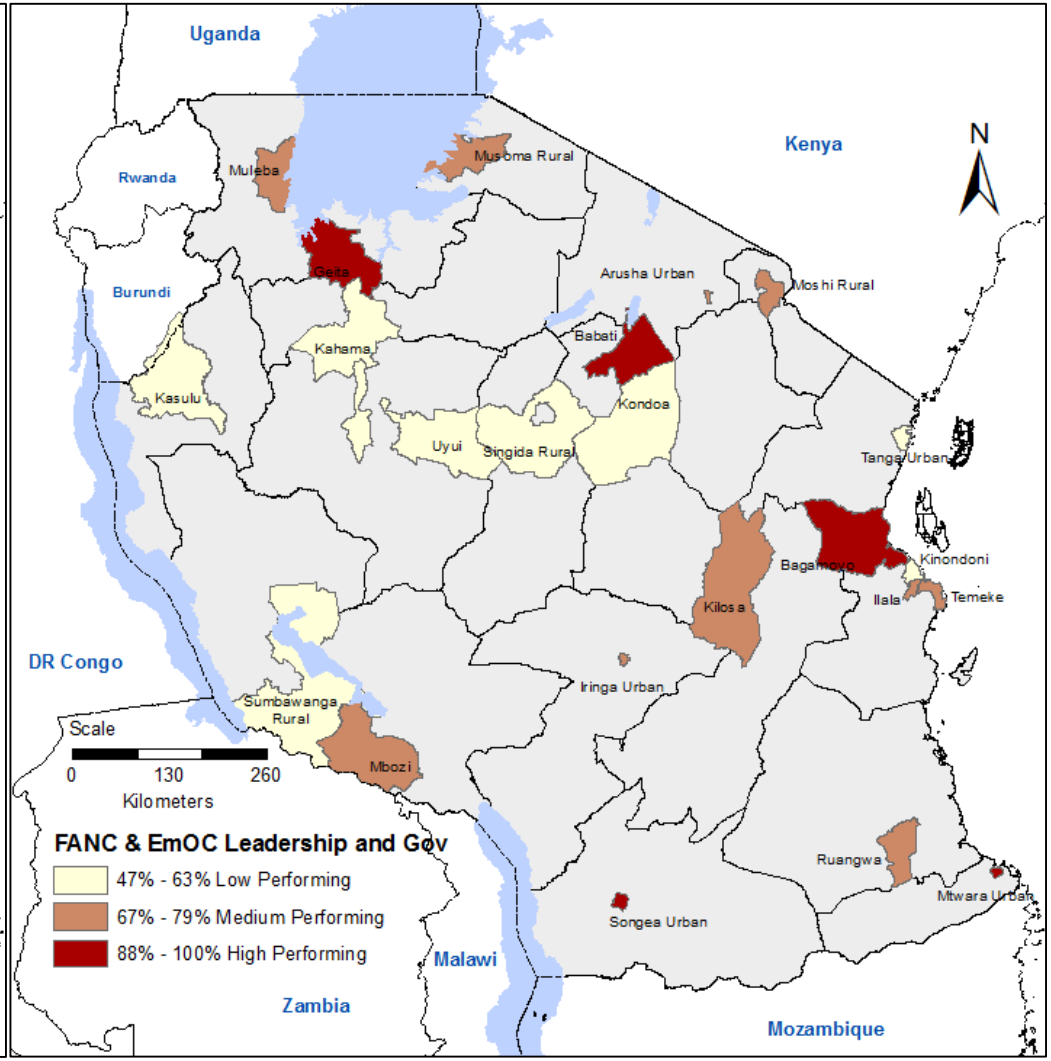


Figure 5.8: Performance on FANC & EmOC Leadership & Governance

Moshi Rural and three urban districts of Ilala, Arusha Urban, and Iringa Urban. All rural districts of Kasulu, Babati, Sumbawanga Rural, and Mbozi were included in both the low performing groups of study districts in stocking FANC and EmOC drugs and supplies. Other districts included in the low performing group on FANC drugs and supplies were Muleba, Kilosa, and Ruangwa whereas other districts in the low performing group on EmOC drugs and supplies were Geita, Musoma Rural, and Bagamoyo. Figure 5.7 shows the implementation strength for the health information system component, with Geita, Kahama, Arusha Urban, Temeke, Ruangwa, and Iringa Urban in the high performing group of districts. These districts scored between 40.1% and 54.5% for the component. Lower performing districts had scores between 18.6% and 23% and included Musoma Rural, Moshi Rural, Kinondoni, Sumbawanga Rural, and Kasulu. All other districts fell into the medium performing group, with scores between 24% and 40%. Figure 5.8 shows performance of study districts on the leadership and governance component. Districts scored a minimum of 47% and a maximum of 100% for the leadership and governance component. The lower performing group of districts included eight districts, whose scores ranged from the minimum score to 63%. The high performing group had five districts scoring between 88% and 100%. Three of the high performing districts were rural districts of Geita, Babati, and Bagamoyo.

Compared to all other five components, study districts had on average the best performance on the leadership and governance component and performed the worst on the human resources component. For each programme component, the high performing group of districts was mostly dominated by urban districts, with rural districts dominating the low performing group. Few rural districts featured in the high performing group and few urban districts featured in the low performing group.

5.4.2 Principal Component Analysis

5.4.2.1 PCA results for FANC programme components

Table 5.4 shows the PCA results from the first and the second principal components for all six programme components and their associated eigenvalues. The two principal components were responsible for over two-thirds (67.4%) variation of the original data with 45.7% and 21.7%for the first and second principal component

respectively. Consecutive principal components had eigenvalues less than 1.0 and were excluded from this analysis.

Table 5.4: FANC Principal Components with eigenvalues >1

FANC Programme Components (variables)	Principal Components	
	1	2
Human resources	0.4893	-0.0342
FANC Service Delivery	0.4819	-0.2315
FANC Drugs and Supplies	0.432	-0.0757
Health Financing	0.5152	-0.1944
Health Information System	0.2103	0.6582
Leadership and Governance	0.1792	0.6845
Eigenvalues (Proportion variation)	2.741 (45.7%)	1.2995 (21.7%)

The first principal component gives high weights to the four of the FANC programme components (original variables): health financing, human resources, FANC service delivery, and FANC drugs and supplies human resources. The first principal component increases with increasing health financing, human resources, FANC service delivery and FANC drugs and supplies human resources The first principal component gives highest weight to health financing – implying that high performing districts (with strong programme implementation) tend to have relatively more funding allocation than low performing districts. The second principal component gives highest weight to leadership and governance and health information system. The negative sign in the loadings implies that the programme component in question correlates negatively with the implementation strength.

5.4.2.2 PCA results for EmOC programme components

Table 5.5 shows the PCA results from the first and second principal components for all six programme components and their associated eigenvalues. The two principal components were responsible for about 71% variation of the original data with 48.2% and 22.7% for the first and second principal component respectively. Other principal components had eigenvalues less than 1 and were excluded from this analysis. The first principal component for EmOC gives high weights to four of the original variables, namely, health financing, EmOC service delivery, human resources and EmOC drugs and supplies. The first principal increases with health financing, EmOC service delivery, human resources, and EmOC drugs and supplies – implying that high performing districts tend to have relatively more funding allocation, higher

Table 5.5: EmOC Principal Components with eigenvalues >1

EmOC Programme Components (variables)	Principal Components	
	1	2
Human resources	0.4837	0.0627
EmOC Service Delivery	0.4842	-0.2011
EmOC Drugs and Supplies	0.4724	-0.1355
Health Financing	0.5186	-0.1129
Health Information System	0.125	0.6863
Leadership and Governance	0.1542	0.6734
Eigenvalues (Proportion variation)	2.893 (48.2%)	1.3627 (22.7%)

EmOC service delivery, more staffed with human resources, and have better stocks of drugs and supplies than low performing districts. The second principal component gives highest weight to the health information system components and leadership and governance. As it was for FANC programme, the second principal component increases with increase in health information system and leadership and governance – implying that high performing districts are associated with health facilities having high reporting and timing rates of HMIS reports and better supportive supervision visits to health facilities.

5.4.3 Sensitivity Analysis

Figures 5.10 and 5.11 show the results from sensitivity analyses of the data for the association between FANC and EmOC programme components and the overall FANC strength index and the overall EmOC strength index respectively. To determine the extent in sensitivity, each component was varied from 25%, 50%, 75%, to 100% of its original weighted score on the overall programme strength. For FANC programme (Figure 5.10), the implementation strength score appears to be particularly sensitive to the four components of human resources, drugs and supplies, health financing, and service delivery. For the EmOC programme (Figure 5.11), implementation strength appeared more sensitive to the human resources, drugs and supplies, service delivery, and health financing components. When component scores were set at 100%, human resources and drugs and supplies exceeded 100% for FANC programme. Likewise, for EmOC programme, variation to 100% in components had the human resources, drugs and supplies, and health financing also exceeding 100%. Sensitivity analyses results seem to confirm results from other analysis (on maternal health experts, programme data from districts, and PCA results) that the overall implementation strength of a programme is most

sensitive to the four components of human resources, service delivery drugs, and supplies, and health financing.

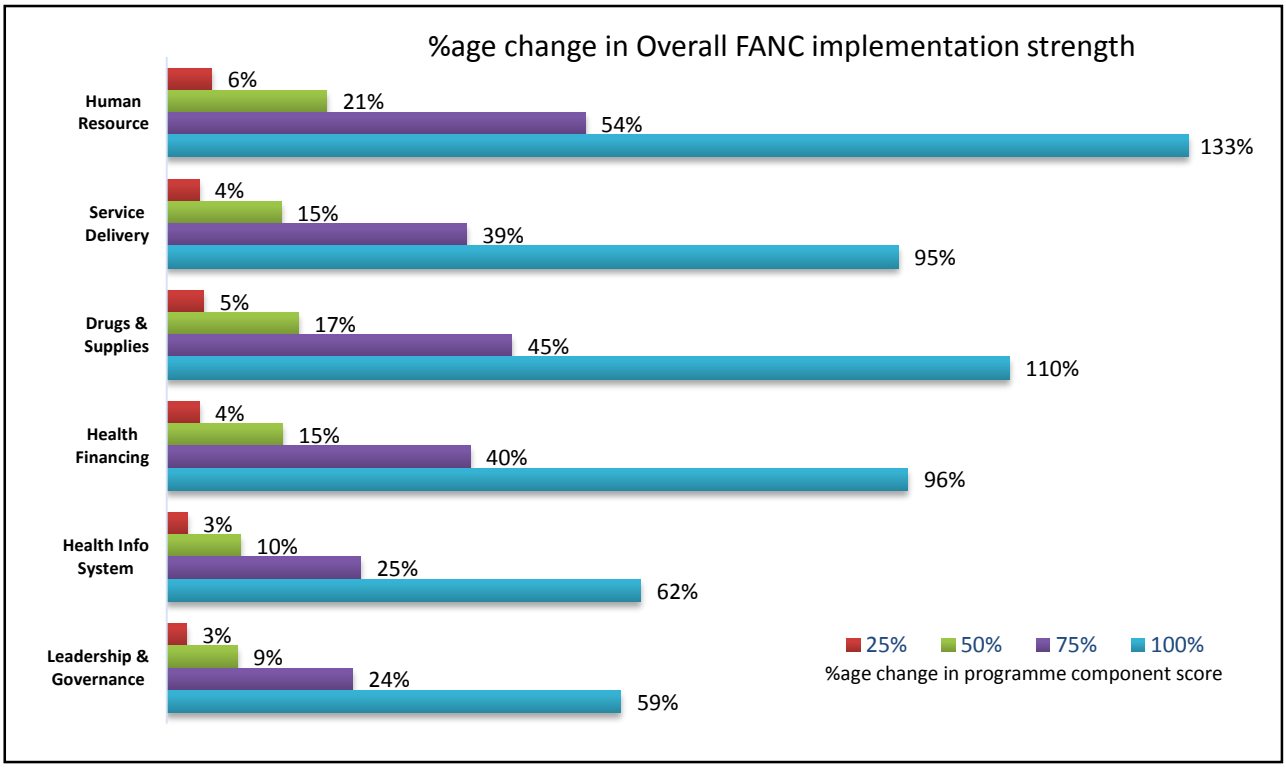


Figure 5.9: Sensitivity analyses for the association between the overall FANC implementation strength and programme components

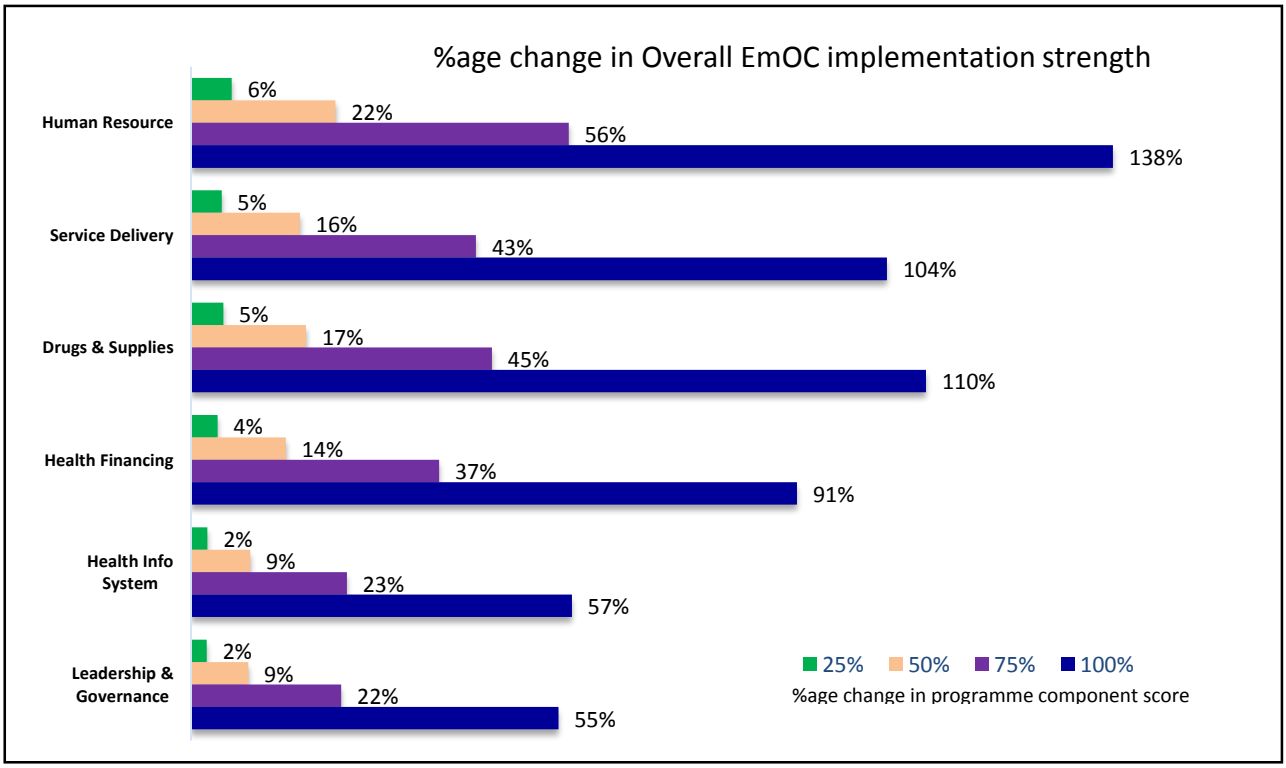


Figure 5.10: Sensitivity analyses for the association between the overall EmOC implementation strength and programme components

5.4.4 Reliability

Tables 5.6 and 5.7 shows the Cronbach's alpha scores for components of the two programmes. The standardised Cronbach's alpha scores summarising the average correlation of components with the overall composite scores, was 0.8271 and 0.8325 for FANC and EmOC programmes respectively, implying that the index has high internal consistency. However, for both programmes, health information system and leadership and governance components had relatively lower scores of correlations on item-test and item-rest compared to other components, suggesting that the two components were less correlated with other components.

Table 5.6: Cronbach's alpha scores for FANC programme components, with overall test scale, standardised to mean=0 and variance=1

Item	Obs	Sign	item-test correlation	item-rest correlation	average inter-item correlation	alpha
Human resources	23	+	0.7833	0.6818	0.3790	0.7855
Service Delivery	23	+	0.7317	0.6122	0.3958	0.7972
Drugs and Supplies	23	+	0.7003	0.5709	0.4061	0.8040
Health Financing	23	+	0.7931	0.6952	0.3758	0.7832
Health Info System	23	+	0.4793	0.2994	0.4784	0.8462
Leadership and Governance	23	+	0.4384	0.2525	0.4917	0.8530
Overall Composite Score	23	+	0.9783	0.9661	0.3152	0.7342
Test scale					0.4060	0.8271

Table 5.7: Cronbach's alpha scores for EmOC programme components, with overall test scale, standardised to mean=0 and variance=1

Item	Obs	Sign	item-test correlation	item-rest correlation	average inter-item correlation	alpha
Human resources	23	+	0.8155	0.7272	0.3791	0.7856
Service Delivery	23	+	0.7512	0.6393	0.4003	0.8002
Drugs and Supplies	23	+	0.7375	0.6212	0.4048	0.8032
Health Financing	23	+	0.8205	0.7342	0.3775	0.7844
Health Info System	23	+	0.3958	0.2062	0.5175	0.8655
Leadership and Governance	23	+	0.4475	0.2644	0.5004	0.8573
Overall Composite Score	23	+	0.9753	0.9615	0.3265	0.7441
Test scale					0.4152	0.8325

5.5 Discussion

The main aim of this chapter was to generate the composite scores of FANC and EmOC programmes that each study district scored according to achievement levels of individual programme components. The chapter showed how the ten steps suggested by the OECD handbook were followed in creating the composite scores.

Important to this chapter was the ranking of study districts and the assessment of the reliability and sensitivity of the composite FANC and EmOC programme indices.

Based on the indicators and the approaches used to generate the composite scores, Arusha Urban district scored the highest on the implementation strength with an overall score of 86% and 80% for FANC and EmOC programmes respectively. Other high scoring districts after Arusha Urban with their corresponding scores for FANC and EmOC respectively, were: Iringa Urban (61%, 57%), and Ilala (56%, 53%). The least performing districts (those with lowest overall scores) for both FANC and EmOC programmes respectively, were: Musoma Rural (24%, 22%), Kasulu (26%, 25%), and Uyui (30%, 25%).

PCA results have shown that all six programme components form 'naturally' into two groups: the first four (human resources, service delivery, drugs and supplies, and health financing – with high loadings in the first principal component) and the last two (health information system and governance and leadership – with high loadings in the second principal component). This grouping suggests that all six components are needed in the programme implementation. Item analysis (Cronbach's) confirmed the PCA results – that all six programme components are useful in contributing to programme implementation strength. However, as the results showed, some programme components contributed more to the overall programme implementation strength than other components. PCA results also supported the opinions of the maternal health experts on the four of the six programme components with the greatest contribution to the overall implementation strength of programmes.

Sensitivity analyses also detected some similar bearings that, the overall implementation strength was most sensitive to the same four components, consistent with the opinions of maternal health experts and the PCA results. The human resources component appeared to be particularly contributing the most strength in both programmes including drugs and supplies, service delivery and health financing. These findings suggest that adequate staffing of doctors, nurses and midwives, sufficient stocking of drugs and supplies in health facilities, health financing and optimal service delivery together play a key role in successful implementation of the two programmes of maternal health services. Even though

reporting of health statistics from health facilities to district levels and provision of regular supportive supervision to health facilities by district health authorities did not seem to contribute much towards the overall implementation strength of programmes, their inclusion in the implementation of FANC and EmOC activities is acknowledged and they are important for the overall implementation strength.

There are several limitations in this sub-study most noticeable of which are the selection of indicators (in terms of number and type), issues with weighting, and the quality of data. It is possible that the type and the number of indicators used in collecting data from study districts have had influence in the results of the implementation strength as well as the ranking of study districts. For example, while other indicators of FANC and EmOC programmes were well reported and had sufficient level of quality (such as indicators for drugs and supplies, health financing, human resources and health information system), measurement of the service delivery component used only three of the seven recommended indicators by the World Health Organisation. It is also possible that the selection of maternal health experts (whose opinions were used to generate the weighting scale) is likely to have influenced the quality of the results. For example, maternal health experts originated from diverse backgrounds with different levels of expertise, different working experience, and from different practice levels of the health system (Chapter 3). It is also possible (as was previously highlighted) that while they had such differences in background, assigning their opinions with equal weights is likely to have contributed to some measurement bias. Even so, the use of expert opinions has been widely reported in the literature – as in the work by Ross and Begala.[189] Rose and Begala used expert respondents in measuring levels and types of programme efforts and estimated the overall maternal and neonatal program effort index in 55 countries.

The quality of data is key to the validity of any research study. While every effort was made in ensuring the quality of data from identified data sources, it is possible that some of the data collected from in study districts could be of questionable quality and could therefore have affected the results of this sub-study. For example, besides district coordinators collecting regular data from health facilities, they could possibly have no control over the authenticity of the data they regularly collected from health facilities for collation at the district level. The quality of health facility-level data lies

with service providers who record service use in numerous registers and record books.

Compared to other indicators, composite indicators can have several advantages and disadvantages. According to Smith et al,[190] composite indicators include the following advantages: they place performance at the centre of the policy arena, they enable judgements to be made on system efficiency, and present the big picture which can be easy to interpret. Smith et al also highlighted the following disadvantages of composite indicators: that, by aggregating individual performance measures, composites may disguise serious failings in some parts of the system (becoming difficult to determine the source of poor performance) and that a composite that is comprehensive in coverage may have to rely on poor quality data in some dimensions.

Tsui et al also pointed out an important shortfall of the composite indicators in that they can be unreliable for comparing "apples and oranges" when constituent indicators have no common unit of measurement.[178] Tsui et al's view point could have been true in this study in incidences for example indicators of the service delivery component and those for the health financing component (district revenues and expenditure) had different units of measurements with no common denominator, for which, straightforward summation of individual component scores may not have been a valid approach and could have incurred "a substantial loss of reliability".[191] To moderate such effects, this study used indicators with common units where possible (such as percentages) and normalised scores by rescaling them to take values between the minimum value and the maximum value.

From time to time, computations of composite scores have involved assigning equal weights to constituent indicators when, in reality, some indicators could be more important to the summary index than others could. For example, the number of qualified staff available at a district hospital could be more important in delivering maternal health services compared to the number of supportive supervision visits conducted by the district health management team. Using statistical analysis (usually correlation coefficients) can be a better way to resolve bias in weighting. Correlation coefficients can be generated using factor analyses or through regression analyses in which coefficients of the predictor variables in the equation generated by the data are taken as weights such as those in the quality of life index of the

Economist Intelligence Unit. Sometimes, weighting issues are resolved by using preference weights generated from surveys. For example in generating an index for quality of life, Hagerty and Land suggested that "if a survey is available to measure the distribution of citizens' importance weight for each indicator, then agreement is maximized when the index is constructed using the mean weights of citizens."[192] In the same way, this study conducted a survey of the maternal health experts from which preference weights were generated and applied to composite indicators of the implementation strength. Imputation of missing data is another likely limitation of any study. For example, the effect of imputing missing values had elevated scores of some of the known poor performing districts on the health information system component. This was observed for the Ruangwa district whose 2011 data had 387 actual HMIS reports. The district did not have data for the number of its timely reported HMIS reports and in its place, the district was imputed with the mean of the rural districts of 376 – resulting in an overall score of 97%, far above the scores of its other indicators.

Despite their observed methodological issues and limitations, scores generated from composite indices have been applied by several global bodies to estimate national, regional, and local metrics on different issues. For example, in 2007, the World Health Organization generated country profiles of environmental burden of disease showing the impact of environmental factors on member states.[193] Likewise, composite scores from the Economist Intelligence Unit's quality of life index have recently been used to identify the best countries in which to be born in 2013.[194] In general, composite indices offer a summarised assessment of a complex problem in a way that the public can easily understand and are therefore more likely to continue to be in used in many policy areas.[106]

5.6 Conclusion

Designing composite indices encompasses technical and analytical issues which can have policy implications. In order to minimise the effect of the highlighted limitations in constructing composite indices and the likely policy implications, analysts should take into account such issues and ensure they are adequately addressed. Based on the minimum set of indicators that can be agreed among national and district stakeholders, composite indicators can have the advantage of using existing data, such as yearly census projections from the national bureau of

statistics, routine data collection by districts and other official data sources generating regular indicators of relevance – thereby minimising the operational cost involving data collection.[195]

District rankings based on implementation strength scores can be used as a policy analysis tool and as an accountability tool for strengthening the district's health system. As a policy analysis tool for maternal health programmes, implementation strength scores can point out the best performing components, from which programme implementers can adopt their technical and logistical delivery approaches leading to efficiency. Improving efficiency in delivering a particular component can potentially have some economic benefits in diverting district resources to other struggling components. Likewise, as an accountability tool for health system strengthening, implementation strength scores can be used by district health care stakeholders (such as funding bodies, constituency councillors and communities) to hold the local government authority to account for low rankings. Future work on this study will benefit if data collection involved carefully selected set of indicators (minimum set) that are considered representative of the programmes. Data collection should also be extended to at least 3 years to be able to perform time series analysis that can potentially give a better picture of the implementation of programmes in study districts.

Chapter 6: Coverage of FANC and EmOC Programmes

6.1 Introduction

The World Health Organisation defines antenatal care as "the care that a woman receives during pregnancy".[196] Antenatal care is an important set of services provided to a pregnant woman, and includes services aimed at promoting health and nutritional support as well as prevention, detection and treatment of anaemia, malaria, sexually transmitted infections and counselling on birth preparedness. While traditional antenatal care used the "risk approach" with which pregnant women are classified according to their likelihood of experiencing complications and promotes frequent visits for better pregnancy outcomes,[197] [198] the alternative focused antenatal care model , is centred around delivering high quality individualised care targeted to needs and limited to four antenatal care visits during pregnancy.[199]

Evidence from several studies since early 2000s has supported the switch to the focused antenatal care model especially in low- and middle-income countries on the grounds that the new model has less burden on the health system and on pregnant women and their families.[200] A well prepared guideline on focused antenatal care provides the following underlying principles for which care should be provided to pregnant women:[201] that care should be woman-friendly, a woman's partner and any other family member should be included during care visits, providers should respect household decision-making, care provided should be culturally appropriate, and individualised (by taking into consideration all of the information known about a woman – current health, medical history, daily habits and lifestyle, household situation, cultural beliefs and customs, and other unique circumstances - the skilled provider can individualise components of care for each woman), as well as care being integrated (to include STI and HIV testing/counselling, malaria detection and prevention, micronutrient provision, birth planning, emergency planning and family planning counselling).

On the other hand, the handbook on monitoring emergency obstetric care,[128] underscored that the signal functions (Table 1.1) are the essential medical interventions for averting life-threatening direct obstetric complications and stresses the importance of health facilities in monitoring the functions as being the essential indicators for monitoring a full range of obstetric services for women with complicated pregnancies. To determine the situation of EmOC in Tanzania, Malecela et al conducted a cross-sectional study for Mainland Tanzania's public health facilities.[202] The situation analysis found that, while all hospitals in Tanzania were supposed to provide comprehensive EmOC services (all 9 signal functions), only over 64% of the government hospitals provided the services. In addition, for the lower health facilities (health centres and dispensaries), the situation analysis also found disappointing results in which, despite delivery services being provided in 81% and 87% of surveyed health centres and dispensaries respectively, only 5.5% of the health centres and none of the dispensaries provided basic EmOC services. Several factors can contribute to such levels of EmOC services provision. For example, Olsen et al conducted a study in northern Tanzania around same period as the situational analysis by Malecela et al and found that poor quality of care in health facilities was the main hindrance of women from accessing EmOC services and neither was their ignorance of the services nor their ability to access the facilities.[203]

Since early 2000s, Tanzania has been scaling up focused antenatal care within the health system with support from various partners including the two pioneering programmes of ACCESS (Access to Clinical and Community Maternal, Neonatal and Women's Health Services) and MNH (Maternal and Newborn Health) both under Jhpiego with funding from the United States of America Agency for International Development (USAID).[204] According to the Tanzanian Demographic and Health Survey (TDHS 2010),[161] most pregnant women (96%) attend at least one antenatal care visit, even though only 43% of them have the recommended four or more visits.[161] In order to scale up coverage of EmOC in Tanzania, sustained and multifaceted efforts to tackle both health-system factors and non-health system factors need to be implemented. These efforts should include (but not limited to):[205] health system strengthening, improving the quality of EmOC services in lower health facilities, and strengthening public–private partnership in the continuum of care. To implement such and more other efforts, the Tanzanian government laid out the 2008 – 2015 plan called "the national road map strategic plan to accelerate reduction of maternal, newborn and child deaths"[206] in which, the MOHSW has used

three strategies in reducing maternal and newborn mortality: policy and advocacy, capacity building, and quality and performance improvement.[207] On policy and advocacy, the government with support from partners developed and disseminated national guidelines on how to provide both FANC and EmOC. On capacity building, the MOHSW and partners have developed educational materials and has continued to provide both pre-service and in-service training, and, on improvement of quality of the services, the MOHSW and partners has and continue to provide interventions on several quality improvement-related factors including "knowledge and skills, motivation, and availability of key resources, supplies and equipment."[207]

The aim of this chapter was to describe the pattern and levels of utilization and coverage of focused antenatal care and emergency obstetric care in the 23 study districts for the period between January 2010 and December 2011 using selected indicators. All data on utilization and coverage of FANC and EmOC programmes were drawn from the central database (DHIS) and the SARA survey – all previously discussed in Sections 2.7.2, 4.2.2, with more discussion presented in Section 6.2.2 below.

6.2 Methods

6.2.1 Study design

Through access to a centralised database, this study reviewed two years' worth of monthly and quarterly records (from January 2010 to December 2011) on utilization and coverage of antenatal and emergency obstetric care from health facilities in 23 districts of Tanzania.

6.2.2 Variables of interest and data sources

6.2.2.1 FANC coverage

The study used three proxy indicators for assessing FANC programme utilisation and coverage (Table 2.4). The first indicator, "antenatal care coverage" was used to assess the programme's coverage and the second and third indicators, "proportion of antenatal care clients testing for HIV/AIDS" and "proportion of antenatal care clients receiving two to five doses of tetanus toxoid injections" were used for

assessing the utilization of antenatal care services in health facilities. The three selected indicators had higher data completion rates (better reported) and were of considerably better data quality than other FANC coverage indicators (Table 2.2).

The first antenatal care visit is a "registration" visit important for initial procedures to assess/prepare a woman for pregnancy and delivery. This includes full history, examination, initial blood tests, immunization, and essential counselling on birth preparedness. The number of clients attending their first antenatal care visit was therefore used as the numerator for the coverage indicator, whose denominator was the number of expected live births in a district per year. The number of expected live births per year was calculated using crude birth rates from national statistics and projections of district populations for the years 2010 and 2011. The Tanzanian National Bureau of Statistics estimated crude birth rates for 2010 and 2011 as 39.3 live births per 1,000 populations and 38.4 live births per 1,000 populations respectively.[208] Number of live births in a year was used as denominator for the other two utilization indicators. Number of antenatal care clients testing for HIV/AIDS and those receiving two to five doses of tetanus toxoid injections were obtained from the DHIS database.

6.2.2.2 EmOC coverage

In order to assess coverage of the EmOC programme, the study used three proxy indicators, all of which are among the eight indicators listed in the United Nations handbook for monitoring emergency obstetric care.[128] The three indicators used (also appearing in Table 2.4) included: the availability of EmOC services, the rate of institutional deliveries, and the rate of Caesarean sections. According to the UN handbook, availability of EmOC services is "measured by the number of facilities that perform the complete set of signal functions in relation to the size of the population".[128] The handbook stresses that a health facility is considered to be 'fully functioning' if in the three months prior to health facility assessment it had conducted the full set of the signal functions of EmOC (Table 1.1). The United Nations recommends that, for every 500,000 population, at least five health facilities should be available to provide basic EmOC services, of which one should provide comprehensive EmOC services.[129] Because presence of health facilities in a district does not automatically qualify the district for the provision of EmOC services, the indicator on availability of EmOC services was centred on assessing the nine signal

functions. District scores on the signal functions were computed using data from the SARA survey. SARA survey was conducted during the same period as this study's data collection and involved an average of 68% of each district's facilities (Section 4.2.2). Survey enumerators interviewed health facility personnel on whether facilities had conducted each of the nine signal functions in the three months prior to the survey.

The second indicator for EmOC programme coverage was "institutional delivery rate". Institutional delivery rate was defined as the proportion of women giving birth in health facilities (hospitals, health centres, or dispensaries) in specified period. Delivering in a health facility increases the chance of a woman being attended by a skilled birth attendant. In Tanzania, skilled birth attendants include physicians, nurses and midwives and exclude traditional birth attendants and maternal and child health aides. The numerator for the indicator was defined as the number of women giving birth in health facilities and its corresponding denominator was the number of expected live births in a district during the same period. The numbers of women giving birth in health facilities were available from the DHIS database.

The third indicator was the rate of Caesarean sections performed in districts. This was defined as the proportion of all births in the population delivered by Caesarean sections in facilities. The numerator was the number of Caesarean sections performed in 2010 and 2011. The denominator was the number of expected live births in each district in the same period (2010 and 2011). The rate of Caesarean section deliveries in a district is an indication of "access to and use of a common obstetric intervention for averting maternal and neonatal deaths and for preventing complications such as obstetric fistula".[128] The UN recommends Caesarean section deliveries be within a range of 5%–15% of all deliveries.[128] Numbers of Caesarean sections performed in districts was reported by health facilities and were available for use in the DHIS database.

6.2.3 Data quality assessment and cleaning

An overview of the overall quality of data for this study was presented in Chapter 2 on Section 2.5 (training, logistics and data collection). Data quality assessment for this study was discussed in Chapter 4 under Section 4.2.3 (data quality assessment and data cleaning). Some of the implementation strength indicators shared the same

data sources with the indicators for programme results. For example, all three indicators of FANC programme results and two of the three indicators of EmOC programme results drew data from the DHIS central database – similar to health information system component indicators. In addition, the third indicator of EmOC programme results (availability of EmOC services) obtained data from SARA survey – same data source to drugs and supplies component of the implementation strength.

After retrieving all relevant data from the DHIS central database, data cleaning was performed before transferring to Stata version 13 for analysis. Hellerstein defines data cleaning as "computational procedures to automatically or semi-automatically identify and, when possible, correct errors in large data sets".[209] Data errors in the DHIS database could have been introduced through reporting from facilities, during data entry, or during integration of data from the multiple sources. For example, the number of deliveries (including Caesarean section deliveries) appeared in three different tables in the DHIS database: in quarterly facility reports, (DHIS table F004-page2); in annual district, attendance reports (DHIS table F005-page2) and in annual district labour and delivery reports, (DHIS table F005-page6).

Data cleaning involved detecting and correcting (by replacing or modifying) or deleting any fraudulent, inaccurate, or irrelevant data points in the dataset[210] – an exercise that involved the author, the two trained DHIS data managers and the reporting district coordinators who were asked to crosscheck with original data sources from health facilities and respective districts. Overall missingness in data was low. For example, for FANC programme results, Moshi Rural, Kilosa, Muleba, and Singida Rural had one or two data points missing in either one or two indicators for either 2010 or 2011 or both, and likewise for EmOC programmes with Kilosa, Sumbawanga Rural and Uyui districts. Rural mean values were used for single imputations in missing data points for the districts because all districts with missing values were rural.

6.2.4 Analytical methods

Health facility data were exported from the DHIS database to Stata version 13, and summarized using tables. For each indicator, summary statistics were generated for all 23 study districts. For uniformity of comparison with other districts, all indicators for FANC and EmOC were expressed as percentages and a summary score was

generated by averaging the percentages of all indicators for each district. Indicators for FANC included expected number of live births per year, percentage coverage of antenatal care services, protection rate of pregnant women against tetanus, and percentage HIV/AIDS testing coverage. Indicators for EmOC included utilization and coverage of emergency obstetric care services (based on 3-month retrospective survey of the nine signal functions), institutional delivery rate and on Caesarean section rate.

For most indicators, a hypothesis test was performed to compare districts based on residence (urban versus rural) to detect any contributing effect on coverage of the FANC and EmOC indicators. For example, with the assumption that urban populations have relatively better access to health services than rural districts, a null hypothesis – H_o: urban districts have similar HIV/AIDS testing coverage as rural districts – was tested using a two-sample Wilcoxon rank-sum (Mann-Whitney) test to show the significance level (P-value) of the differences observed between urban (n=8) and rural (n=15) districts.

6.3 Results

6.3.1 Coverage and utilization of antenatal care

Table 6.1 presents results on both utilization and coverage of antenatal care for the FANC programme. Results on the three indicators used are presented and discussed below.

6.3.1.1 Antenatal care coverage

Study districts had high coverage of antenatal care services with an overall mean of 113%[j], 95% CI (95% 132%), median of 115% for 2010 and overall mean of 106% , 95% CI (86% 126%), median of 113% for 2011. Individual districts, however, had wide-ranging antenatal care coverage, from as low as 39% in 2010 and 35% in 2011 to as

[j] Aggregate rates over 100% could mean numerator was larger than the denominator – for example, more pregnant women attending antenatal care clinics than the anticipated annual number of live births pre-set by Council Health Management Teams. It is also possible that service providers could have included repeat visits in compiling reports to the districts.

Table 6.1: Coverage of antenatal care and utilization of FANC services in study districts for 2010 and 2011

No	Study district	Residence	Expected number of live births		Percentage antenatal care coverage		Percentage tetanus toxoid utilization coverage		Percentage HIV/AIDS testing coverage		District Score
			2010	2011	2010	2011	2010	2011	2010	2011	
1	Arusha Urban	Urban	14,246	14,312	174%	177%	144%	71%	126%	105%	133%
2	Ilala	Urban	28,376	28,036	213%	218%	106%	107%	107%	129%	147%
3	Iringa Urban	Urban	5,999	6,090	100%	116%	45%	56%	92%	72%	80%
4	Kinondoni	Urban	48,620	48,009	113%	117%	57%	24%	78%	31%	70%
5	Mtwara Urban	Urban	4,818	4,874	51%	48%	59%	65%	78%	84%	64%
6	Songea Urban	Urban	6,903	6,995	201%	122%	98%	57%	142%	87%	118%
7	Tanga Urban	Urban	11,746	11,739	130%	135%	67%	21%	70%	33%	76%
8	Temeke	Urban	34,230	33,793	136%	172%	95%	81%	103%	67%	109%
9	Babati	Rural	16,026	16,258	75%	74%	50%	41%	59%	64%	61%
10	Bagamoyo	Rural	10,913	10,897	91%	80%	60%	37%	79%	65%	69%
11	Geita	Rural	33,644	33,786	134%	146%	62%	70%	54%	27%	82%
12	Kahama	Rural	32,036	32,591	122%	104%	87%	67%	44%	23%	75%
13	Kasulu	Rural	24,811	25,108	92%	119%	58%	78%	17%	55%	70%
14	Kilosa	Rural	23,107	23,056	115%	35%	74%	17%	58%	37%	56%
15	Kondoa	Rural	19,627	19,521	83%	82%	54%	58%	60%	14%	58%
16	Mbozi	Rural	26,028	26,188	123%	113%	57%	40%	65%	82%	80%
17	Moshi Rural	Rural	18,160	18,032	39%	46%	34%	67%	19%	48%	42%
18	Muleba	Rural	18,697	18,908	127%	130%	72%	70%	45%	46%	82%
19	Musoma Rural	Rural	16,505	25,704	67%	50%	92%	63%	59%	35%	61%
20	Ruangwa	Rural	5,777	5,746	76%	68%	65%	64%	62%	61%	66%
21	Singida Rural	Rural	19,135	19,106	87%	70%	44%	37%	54%	45%	56%
22	Sumbawanga Rural	Rural	18,981	19,184	143%	103%	125%	67%	50%	10%	83%
23	Uyui	Rural	13,856	13,925	117%	118%	103%	106%	92%	61%	99%
Urban Districts				Mean (Median)	140% (133%)	138% (128%)	84% (81%)	60% (61%)	100% (97%)	76% (78%)	100% (95%)
Rural Districts				Mean (Median)	99% (92%)	89% (82%)	69% (62%)	59% (64%)	54% (58%)	45% (46%)	67% (69%)
Overall				Mean (95% CI) (Median)	113% (95% 132%) (115%)	106% (86% 126%)	74% (63% 86%)	59% (48% 70%) (64%)	70% (57% 85%)	56% (37% 67%) (55%)	80% (68% 90%) (75%)
2-sample Wilcoxon rank-sum test for H_0; urban=rural				P-value	0.061	0.02	0.302	0.838	<0.001	0.01	0.0201

Source: DHIS database

high as 213% in 2010 and 218% in 2011. For both years of assessment, urban districts had higher antenatal care coverage than rural districts, with 2010 means (medians) of urban=140% (133%) versus rural=99% (92%) and with 2011 means (medians) of urban=138% (128%) versus rural=89% (82%). The differences in antenatal care coverage between urban and rural districts are significant for 2011, with P–value=0.02; but of borderline significance for 2010 (P–value=0.06). The top three districts with high antenatal care coverage for 2010 and 2011 were all urban districts. These were, Ilala, Songea Urban and Arusha Urban (in 2010) and Ilala, Arusha Urban and Temeke (in 2011). Moshi Rural, Mtwara Urban, and Musoma Rural were the lowest performing districts on antenatal care coverage in 2010. In 2011, the lowest performing districts were Kilosa, Moshi Rural, and Mtwara Urban.

6.3.1.2 Tetanus toxoid coverage rate

On average, about three quarters of pregnant women attending for their first visits in antenatal care clinics received between 2 to 5 injections of tetanus toxoid in 2010; mean=74% (95% CI: 63%–86%). In 2011, this dropped to 59% (95% CI: 48%–70%). For both years, the proportion of pregnant women in urban districts who received 2–5 doses of tetanus toxoid was higher than that in rural districts. In 2010, a mean (median) of 84% (81%) of pregnant women in urban districts received 2–5 doses of tetanus toxoid injection compared to 69% (62%) in rural districts. Similarly in 2011, a mean (median) of 60% (61%) of antenatal care clients in urban districts received 2–5 doses of tetanus toxoid injection compared to 58% (63%) in rural districts. Overall, the differences between urban and rural districts in coverage of tetanus toxoid vaccination for both 2010 and 2011 appeared not significant, P-value=0.302 (for 2010) and P-value=0.838 (2011). Musoma Rural, Mtwara Urban, Moshi Rural, Sumbawanga Rural, and Uyui were the districts with highest coverage of tetanus toxoid vaccination in 2010. In 2011, the leading districts were Mtwara Urban, Musoma Rural, and Ruangwa. Districts with the lowest coverage of tetanus toxoid vaccination were Iringa Urban, Mbozi and Geita in 2010 and Tanga Urban, Kinondoni and Mbozi in 2011.

6.3.1.3 HIV/AIDS testing rate

On HIV/AIDS testing coverage in study districts, the overall mean (95% CI) for 2010 was 71% (57%–85%). However, in 2011 HIV/AIDS testing coverage dropped to 52%

(37%–67%). Testing coverage for HIV/AIDS among pregnant women was higher in urban districts for both 2010 and 2011 with mean (median) of 100% (97%), compared to 54% (58%) in rural districts. There is strong evidence for this difference between urban districts and rural districts in coverage of HIV/AIDS testing among pregnant women (P-value<0.001 for 2010 and P-value=0.01 for 2011). The best performing districts in testing of HIV/AIDS among pregnant women for 2010 were Mtwara Urban, Iringa Urban, and Musoma Rural. The best performing districts in 2011 were Mtwara Urban, Ruangwa, and Babati. The worst performing districts in HIV/AIDS testing among pregnant women were Kasulu, Muleba, and Sumbawanga Rural in 2010; and Sumbawanga Rural, Kondoa and Geita in 2011.

The summary index (overall district score) indicated that Ilala district had better antenatal coverage and utilization of FANC services, with a score of 147%. Moshi Rural district tailed other districts, with an overall score of 35%. On average, urban districts had better antenatal care service coverage and utilization, with an overall score of 100% (median=95%) compared to an overall score of 67% in rural districts (median=69%). Difference in the coverage and utilization of the services between urban and rural districts was significant (P-value=0.02).

6.3.2 Coverage of EmOC services

Utilization and coverage of emergency obstetric care services was evaluated based on the availability of EmOC services in study districts (based on scores on the nine signal functions), on institutional delivery rates and on Caesarean section rates (based on actual Caesarean sections conducted in districts). Overall, study districts had an average score of 64% across all three indicators (Table 6.2). Urban districts outperformed rural districts, with an average score of 87% compared to 51%. Best performing districts on overall coverage of EmOC services were all urban districts: Songea Urban (103%), Iringa Urban (102%), Ilala (97%), and Arusha Urban (92%). Worst performing districts were all rural districts: Musoma Rural (31%), Singida Rural (38%), and Geita (43%).

6.3.2.1 Availability of EmOC services in districts

Table 6.2 shows that, in 2011, study districts had 1,706 eligible health facilities that were capable of providing basic EmOC services. These included 1,451 dispensaries,

177 health centres, and 78 hospitals. Availability of EmOC services was not evaluated based on the number and type of health facilities that reported offering the services but on a survey of the nine signal functions. Table 6.2 shows the results of the survey. Overall, study districts had an average score of 65% for provision of the nine signal functions, ranging from as low as 41% (Geita district) to as high as 84% (Arusha Urban district). Again, urban districts had better overall coverage of EmOC services than rural districts, with average scores of 71% and 61% respectively.

6.3.2.2 Institutional delivery rates

There was an overall average score of 60% for institutional delivery, with populations in urban districts delivering in health facilities at a rate of about 2.5 times that of populations in rural districts (urban districts=98%, rural districts=40%). The three districts of Songea Urban, Iringa Urban, and Ilala exceeded their projected statistics on institutional delivery, scoring 136%, 131%, and 117% respectively. Districts with the lowest institutional delivery rates were those of Uyui (22%), Moshi Rural (23%), Musoma Rural (27%), and Singida Rural (29%).

6.3.2.3 Caesarean section rates

Caesarean section rates reflected the number of Caesarean section deliveries performed in study districts as a proportion of the number of live births. To transform Caesarean section rates to a scale between 0% and 100% for comparability with scales in other indicators, all districts with Caesarean section rates below 5% were assigned with a score equivalent to the respective rate divided by 5% (five per cent is the minimum rate required by the World Health Organization for Caesarean sections). Districts with rates of 5% or more were assigned a score of 100%. Table 6.2 shows that seven of the eight urban districts had a 100% Caesarean section rate score, followed by Muleba district (86%) and Babati district (79%). The majority of study districts had conducted the WHO-recommended proportion of Caesarean sections of between 5% and 15% of all live births. Thirteen districts (over half of the

Table 6.2: EmOC service coverage in study districts (with 2010 and 2011 routine data and 2012 survey signal functions)

No	Study District	1) Availability of EmOC Service (Signal Functions Scores)										# HFs	Av LB	Av HFD	Av CS	2) Av IDR	CSR	3) CSR	District Score
		AB	UT	AC	AVD	MRP	RRP	NR	CS	BT	Avrg								
1	Arusha Urban	86%	93%	79%	100%	100%	100%	100%	50%	50%	84%	62	14,279	13,277	2,908	93%	20%	100%	92%
2	Ilala	91%	88%	79%	91%	88%	88%	91%	21%	18%	73%	175	28,206	32,993	2,710	117%	10%	100%	97%
3	Iringa Urban	100%	91%	80%	100%	91%	91%	91%	18%	9%	75%	28	6,045	7,910	1,453	131%	24%	100%	102%
4	Kinondoni	90%	67%	54%	68%	67%	67%	59%	42%	39%	61%	190	48,314	37,249	4,329	77%	9%	100%	79%
5	Mtwara Urban	70%	60%	56%	67%	50%	50%	56%	11%	11%	48%	20	4,846	4,382	727	90%	15%	100%	79%
6	Songea Urban	100%	91%	100%	100%	100%	100%	45%	9%	9%	73%	31	6,949	9,429	1,464	136%	21%	100%	103%
7	Tanga Urban	100%	91%	91%	96%	96%	96%	91%	30%	32%	80%	56	11,743	7,297	889	62%	8%	100%	81%
8	Temeke	74%	84%	74%	89%	100%	100%	95%	21%	21%	73%	148	34,012	25,219	725	74%	2%	43%	63%
9	Babati	100%	100%	97%	100%	92%	92%	83%	11%	8%	76%	45	16,142	5,214	638	32%	4%	79%	62%
10	Bagamoyo	96%	75%	81%	100%	88%	88%	84%	4%	2%	69%	66	10,905	4,387	277	40%	3%	51%	53%
11	Geita	70%	19%	16%	96%	35%	35%	87%	6%	2%	41%	73	33,715	17,719	596	53%	2%	35%	43%
12	Kahama	96%	100%	66%	78%	86%	86%	80%	4%	6%	67%	71	32,314	12,060	883	37%	3%	55%	53%
13	Kasulu	87%	84%	42%	81%	77%	77%	52%	6%	6%	57%	81	24,959	12,934	821	52%	3%	66%	58%
14	Kilosa	100%	96%	98%	98%	100%	100%	100%	2%	2%	77%	77	23,082	6,892	514	30%	3%	52%	53%
15	Kondoa	96%	75%	10%	100%	92%	92%	8%	2%	2%	53%	73	19,574	8,048	608	41%	3%	62%	52%
16	Mbozi	63%	43%	46%	84%	68%	68%	33%	4%	4%	46%	77	26,108	12,384	528	47%	2%	40%	45%
17	Moshi Rural	100%	100%	58%	97%	77%	77%	82%	11%	11%	68%	75	18,096	4,092	460	23%	3%	51%	47%
18	Muleba	84%	100%	22%	97%	100%	100%	81%	11%	11%	67%	42	18,802	11,910	813	63%	4%	86%	72%
19	Musoma Rural	98%	98%	34%	90%	85%	85%	33%	2%	2%	59%	63	21,104	5,642	73	27%	0%	7%	31%
20	Ruangwa	67%	70%	22%	89%	100%	100%	81%	4%	4%	60%	33	5,761	2,860	168	50%	3%	58%	56%
21	Singida Rural	100%	100%	93%	100%	15%	15%	38%	5%	5%	52%	63	19,121	5,461	307	29%	2%	32%	38%
22	Sumbawanga Rural	96%	72%	14%	91%	81%	81%	85%	2%	6%	59%	118	19,083	10,822	514	57%	3%	52%	56%
23	Uyui	100%	100%	100%	100%	61%	61%	92%	0%	0%	68%	39	13,890	3,062	514	22%	3%-	52%	47%
Urban Districts, Mean		89%	83%	76%	89%	86%	86%	78%	25%	24%	71%	89	19,299	17,219	1,901	98%	14%	93%	87%
Rural Districts, Mean		90%	82%	53%	93%	77%	77%	68%	5%	5%	61%	66	20,177	8,232	411	40%	3%	52%	51%
Overall, Mean		90%	83%	61%	92%	80%	80%	72%	12%	11%	65%	74	19,872	11,358	929	60%	6%	66%	64%

AB= parenteral administration of antibiotics, UT=administration of uterotonic drugs, AC=parenteral administration of anticonvulsants, AVD= assisted vaginal delivery, MRP=manual removal of the placenta, RRP=removal of retained products, NR=basic neonatal resuscitation, CS= Caesarean Section, BT=blood transfusion, HFs=Health facilities; Av LB=Average number of live births; Av HFD=Average number of Health Facility Deliveries; AV CS=Average number of Caesarean sections; Av IDR=Average Institutional Delivery Rates; CSR= Caesarean Section Rate;

study districts) had less than 5% of births performed by Caesarean sections between 2010 and 2011. While Kilosa district had submitted no data on Caesarean sections, Sumbawanga Rural and Uyui districts did not perform any Caesarean section deliveries for having no qualified health facilities, although Sumbawanga Rural district started a few operations after the assessment, in 2012.

6.4 Discussion

This study explored the pattern and levels of utilization of specific FANC and EmOC services in study districts. The study also assessed the coverage of the two programmes in the districts by looking at the number of clients receiving the services from health facilities as a proportion of those that would be expected – derived from population estimates. Study data were drawn from two years of 2010 and 2011 and from more than a quarter of health facilities of Mainland Tanzania based in the 23 nationally representative districts.

Coverage rates for antenatal care reported by this study are largely in agreement with those reported by the 2010 Tanzania Demographic Health Survey (TDHS 2010).[161] Using the same definition of coverage (pregnant women having received antenatal care at least once), TDHS 2010 reported that, for Mainland Tanzania, antenatal care coverage was 96% – with 100% coverage in two regions of Kilimanjaro and Dar es Salaam. All three districts of Dar es Salaam region were part of the 23 districts of this study. Both the TDHS 2010 and this study agree on the fact that urban districts had higher antenatal care coverage than rural districts. This is likely due to reasons that in the former there are more health facilities offering antenatal care service than in the latter. Other studies also reported similar findings.[211] [212] This correlation could possibly explain the higher HIV/AIDS testing rates in urban districts than in rural districts, with 66% versus 46% in 2010 and 46% versus 36% in 2011.

The rates of tetanus toxoid vaccination reported by this study were higher than those reported by the TDHS 2010. TDHS 2010 reported that, for Mainland Tanzania, the tetanus toxoid vaccination rate was 48% which are in contrast to the vaccination rate of 69% for2010 reported by this study. While the TDHS reported urban populations had higher rates of tetanus toxoid vaccination than rural populations (61% versus 44%), this study found contradicting results showing that the proportion of pregnant

women in rural districts receiving 2–5 doses of tetanus toxoid was slightly higher than those in urban districts (64% in rural versus 59% in urban for 2010, and 58% in rural versus 39% in urban for 2011). Findings of this study on availability of EmOC health facilities differed from those reported by studies conducted in some sub-Saharan countries. For example, Pearson and Shoo[213] found that, for a population of 500,000, the number of facilities providing basic and comprehensive EmOC were respectively: Kenya (1.1 and 1.6), Rwanda (0 and 4.3), South Sudan (0.5 and 1), and Uganda (1.0 and 1.2). Fauveau[214] also found that, in a survey conducted in Guinea-Bissau, there were only 11 of the required 13 basic EmOC facilities and 3 of the required 5 comprehensive EmOC facilities in a population of over 1.1 million. It is possible that the proportion of health facilities providing basic and comprehensive EmOC services in study districts are likely to be inflated because the SARA survey found out that not all facilities visited had actually conducted the signal functions for either basic or comprehensive EmOC service.

There were some unexpected results from study districts, especially on indicators using the denominator of number of expected annual live births. For example, the three urban districts of Ilala, Iringa Urban, and Songea Urban exceeded their annual expected institutional delivery rates. Districts with institutional delivery rates of over 100% were likely to be associated with delivering more pregnant women from neighbouring districts or had wrong denominators. Unexpected results were also found in Moshi Rural district for the institutional delivery rate. Table 6.2 showed that despite the district having a relatively higher number of health facilities (75) compared to most other study districts, it had one of the lowest delivery rates, at 23%. There are two possibilities for such a finding the most plausible of which could be that Moshi Rural women prefer delivering in the neighbouring Moshi Urban district or that women in Moshi rural prefer to deliver at home or that the denominator for the district was set higher than the reality. Additionally, reporting numbers from health facilities to districts or from districts to the central database could be another area where such errors could have being introduced. Other unexpected findings are those on the overall decline of figures from 2010 to 2011. It is unclear at this point as to what could have contributed to most indicators reporting lower figures in 2011 than in 2010. One likely explanation could be logistical in that SPD started its operations in 2010, it is possible that reporting of numbers could have been improved in 2011 as district coordinators collated figures from study districts with improved quality control of data.

6.5 Limitations

While the study found high coverage and utilization of FANC and EmOC services in study districts, it did not conduct a qualitative assessment (in-depth interviews and/or focus group discussions) to complement findings from women using the services. The study used health facility data which depended on how service providers in health facilities were attentive to its quality. It is possible that, using the National Bureau of Statistics' projections to calculate the number of expected live births could have introduced some unrealistic crude birth rates in study districts. It is possible that women travel to different health facilities for services, and may not go to the one closest or that some districts could have higher estimates and others lower estimates for indicators whose denominator on the number of expected live births was calculated using projections. Better estimates could be obtained through active monitoring of live births using demographic surveillance systems.

For good quality of antenatal care to pregnant women, WHO recommends at least four antenatal care visits with the first visit well timed (between 12 and 16 weeks).[162] The study restricted the numerator for antenatal care coverage to the number of women who attended first visits only. While it was desirable to have the data on number of clients attending four or more visits, such data were either not available (not reported) or were unreliable due to recording issues for most study districts. Antenatal care coverage is however a widely recognized indicator as an important determinant of maternal health care coverage, and is used together with institutional delivery rate for monitoring MDG-5 progress.[215] While data on the rate of Caesarean sections can be collected in population surveys such as demographic and health surveys, data for it indicator (Caesarean section rates) were collected from hospital records as rates based on service statistics. Because Caesarean sections can only be conducted in health facilities, data for this study were therefore considered more precise than population-based rates. Facility data are collected routinely from operating theatre logbooks, which were often the most complete records available in study districts.

It is possible that not all hospitals and health centres provide comprehensive EmOC services and that not all dispensaries provide basic EmOC services. This is substantiated by the findings of the 2006 Tanzania Service Provision Assessment

Survey, which reported that two of the 25 assessed hospitals and as many as 48 of the 55 health centres did not perform Caesarean sections.[216] The study did not extend the assessment to other indicators of FANC services such as proportion of clients receiving the required prophylaxis against malaria in pregnancy, iron deficiency, and other services, but used indicators on vaccination with tetanus toxoid and screening for HIV. The two indicators were adequately reported. It is possible, therefore, that unreported indicators could have provided a different picture on FANC service utilization from the picture portrayed by the two indicators.

According to the DHIS database, the 23 study districts have over 1,700 health facilities, including 1,451 dispensaries, 177 health centres, and 78 hospitals. Most health facilities are publicly owned and some are owned by either nongovernmental (not-for-profit) organizations or the private sector. The number of health facilities in districts fluctuates each year due to closure of some and opening of a few new facilities in some districts. Not all facilities provide reproductive and child health services. Most of the reported data on utilization of FANC and EmOC services originated from HMIS Register Book 6, which records service utilization of antenatal care in RCH clinics, and from HMIS Register Book 12, which records services provided in Labour and Delivery wards.

The overall score of a district was based on either programme's aggregation of its constituent indicators. While assessing FANC programme coverage only three (of potentially many) indicators were used: antenatal care coverage, utilisation of tetanus toxoid in pregnant women, and HIV/AIDS testing rates among antenatal care clients, and, only three of the eight WHO-recommended indicators were used to assess EmOC programme coverage (availability of EmOC services, institutional delivery rates, and Caesarean section rates). It is therefore possible that by including other indicators for both programmes, coverage results would have shown a different pattern and probably presented a much better balanced coverage picture. However, excluded indicators for both FANC and EmOC programmes did not have data reported from health facilities, or its data were sporadic and/or of insufficient quality. Assessment of districts on FANC programme coverage did not include indicators from all five themes of antenatal care service: history-taking, physical examination, laboratory examinations, drug administration and immunization and health education. The five indicators excluded from assessing EmOC programme coverage included met need for EmOC, direct obstetric case fatality rate,

intrapartum, and very early neonatal death rate, proportion of maternal deaths due to indirect causes, and proportion of all births in EmOC facilities.

There may have been some potential for some degree of circularity between the indicators of service availability for the implementation strength (Chapter 4) and the EmOC signal functions (presented in this Chapter) for being closely related. The service availability component of the implementation strength was measured by three indicators of number of facilities per 10,000 population, number of hospital beds per 10,000 population, and the number of first ANC visits per 10,000 population. EmOC programme coverage on the other hand, was measured by percentage scores of the EmOC signal functions. While the higher the number of facilities in study districts may not imply higher availability EmOC services, there is an increased possibility that health facilities having conducted the EmOC signal functions within three months of the SARA survey could be closely dependent on the number of health facilities in study districts.

Another potential limitation with regard to EmOC is the extent to which the same data might contribute both to implementation strength and to the outcome measure of coverage. For EmOC, the implementation strength measure included drugs and supplies on the day of the survey, whereas EmOC outcome measures included provision of services over the previous 3 months. For example, one of the EmOC signal functions is administration of uterotonic drugs in the previous 3 months, which would be impossible if those drugs were never in stock. Although the data source for drug stocks and for administration of drugs in the previous 3 months are the same survey, separate questions were asked on these two issues. I therefore have some confidence that there is a degree of independence of the two measures.

6.6 Conclusion

The high coverage of first antenatal care visits in Tanzania gives an opportunity for reaching almost all women. Although less than half of women attend four or more antenatal care visits, health care providers can still utilise the opportunity of first visits by ensuring pregnant women are provided with all essential services including promotional counselling services to improve repeat visits.

Coverage of emergency obstetric care services is still low. Given that the Tanzanian government has national guidelines for obstetric care service provision, district health authorities should prioritise this area by ensuring essential drugs and supplies are available in health facilities and providing on-the-job training to clinicians and nurses for improved service provision. As expected, urban districts had better statistics than rural areas. Action to reduce this inequity is urgently needed. Tackling such challenges requires long term planning and commitment by the central government and local authorities.

There is also a need for policy formulators, health planners and implementers of FANC and EmOC programmes to develop an agreeable minimum set of essential indicators to facilitate in monitoring and evaluation of the programmes. While the Ministry of Health and Social Welfare has come up with the scorecard comprising of 13 indicators for reproductive, maternal, newborn, and child health (further discussion of the scorecard in Chapter 10),[217] it is important that the list of indicators be inclusive of both the implementation strength and programme coverage as well as some on equity to be able to have a balanced representation of the efforts directed to and the results generated by programmes. Reaching at an agreeable set of essential indicators should be a process that will involve top-down and bottom-up discussions to allow mutual involvement of the parties along the chain of programme implementation. Based on the final list of the agreed essential indicators, planning and implementation of FANC and EmOC services at district and facility levels can be achieved by prioritising efforts (funding and activities) aimed at improving the indicators. For cost-effectiveness, continuous reporting of the essential indicators should be aimed at obtaining data from routine systems that are already established and whose collection is part of the quarterly data reporting. While data quality for all data from routine systems is of priority, the quality of data for the essential indicators should be given higher priority and improved in order to improve evidence-based planning and programme implementation.

Chapter 7: Contextual factors relevant to FANC and EmOC

7.1 Introduction

In analysing associations between maternal health interventions and their effects in low- and middle-income countries, numerous studies have assessed contextual factors.[136 218] Contextual factors can confound the association between an exposure and an outcome variable thereby provide alternative explanations of the effect. For example, a non-governmental organisation providing some form of community-based intervention to empower women socioeconomically can be associated with the implementation a FANC programme (the exposure) in a certain district and at the same time be associated with the increase in number of women delivering in health facilities in the same district (the outcome), the community-based intervention will be a confounder to the association between the exposure and the outcome. Contextual factors can also play a role of effect modification (mediation) by facilitating (positively) or inhibiting (negatively) the results of a certain programme. For example, if the number of women delivering in health facilities somehow differs from one specific age group of the women to another specific age group, then age will be the effect modification of the association between the FANC programme and delivering in health facilities.

Because of the likely role of contextual factors in programme coverage, it is imperative to document and analyse their contribution in the effort to scale programmes in target populations. Mangham and Hanson[219] pointed out four major issues involved with scaling up a health intervention: service delivery issues, cost issues, equity and quality issues, and other scale-up constraints. While most issues can be resolved during planning and during implementation, other factors external to designed interventions are likely to constrain its scale-up. If one of the goals of programme is to scale it to its target population, implementers will benefit more if they gain an understanding of factors in the context of the programme. It is in the interest of funding bodies and local governments to scale up maternal health programmes in order to accelerate achievement of the millennium development goal

(MDG) 5. It is equally important to investigate, document and analyse the levels of factors in context. If contextual factors are not properly documented, interventions may not recognize the magnitude of their effect in delivering maternal health interventions and may not sufficiently associate the resulting effect with programme activities.

This chapter describes the contextual factors associated with implementation of FANC and EmOC programmes in Tanzanian districts. Contextual factors are included in this study as key variables that are likely to explain differences in coverage of maternal health programmes in districts. Contextual factors presented in this chapter include total fertility rate, infant and under-five mortality rates, female literacy levels, water and sanitation, transport and communication, natural disasters and other projects and initiatives operating in a district. Results presented here will be used in the next chapter on illustrative dose-response analysis.

7.2 Methods

7.2.1 Selection of indicators

A review of currently published articles on contextual factors in public health studies was conducted. The search was limited to the PubMed database and included the following search terms: 'contextual factors', 'contextual characteristics', 'influencing factors', and 'moderating factors'. Studies were included if they were published as a free full text article and if they were conducted in a sub-Saharan African country. Further criteria included articles published in English language and if they were available in print or online and if were published between January 2003 and December 2013. Additional list of targeted contextual factors related to implementation of large-scale health programmes were included from the literature.[66 220] Resulting factors were further reviewed based on their relevance to implementation, utilization, and scaling-up of FANC/EmOC programmes and the possibility of getting related data from study districts during the limited time of data collection. The summarised contextual factors were those other than FANC and EmOC programme components that might have affected maternal health in the study districts. Emphasis was directed at contextual factors that might have changed over the two-year study period (January 2010 to December 2011). For each identified contextual factor, a proxy indicator was developed for use in assessing its

status in districts. The final list of factors appears in Table 2.5 and were grouped into six main categories: demographic, socioeconomic, environmental, infrastructural, natural disasters and public health system.

7.2.2 Study design

The study involved a cross-sectional survey of district officials within the District Executive Director's office of each study district. The study also conducted a review of documents and accessed data sources containing relevant information on selected district- and/or national-level indicators.

7.2.3 The questionnaire and respondents

Using the final list of contextual factors and their corresponding indicators, a two-page questionnaire was designed (Appendix 2: Data Collection Form for Contextual Factors in a District). The first draft was prepared and distributed to three (of 23) data collectors who consulted for opinions with district sources and used their own experience in districts to advise on some key content and design issues. Improvements were made and the final version of the questionnaire was updated. The questionnaire contained two modules: a module on 'health' designed to obtain data on projects and initiatives (other than FANC and EmOC) within the public health system that operated in the district. The second module of the questionnaire was designed to elicit data on contextual factors for water, sanitation, transport, communication, and natural disasters. The main respondents to the questionnaire were the district officials of different departments within the District Executive Officer's office. District Executive Officers are the government executive officers at the district level who implement government development programmes and are the official record keepers of all development-related activities in districts.

7.2.4 Other programmes and initiatives

The 'health' module of the questionnaire was administered for all district medical officers – including district health secretaries and the district social welfare officer. Health module respondents were asked to state the names of projects/initiatives in operation within the district and within the period from January 2010 to December 2011. In addition, respondents were asked to name the purposes, areas of coverage

and funding sources for the stated projects/initiatives. Data collectors probed respondents on all projects/initiatives providing services on maternal and newborn health (other than FANC and EmOC programmes), on child health programmes (immunization, micronutrients, nutrition), on malaria projects/programmes (indoor residual spraying, bed net use), on HIV/AIDS projects/programmes (care and treatment, counselling and testing), and on community health/social protection programmes. Data collectors also reviewed district medical officers' reports for further information. The data collectors were the 23 FBIS coordinators who work under the district medical officer and who were themselves familiar with health-related projects within the district. Data collectors also visited departments within the District Executive Director's office to collect data on other indicators.

7.2.5 Water and sanitation

Data on the percentage of households using drinking water from improved water sources were obtained from the water department and/or from the district planning office. Clarification was provided on the term 'improved drinking water sources', which included: piped water into dwelling, piped water to yard/plot, public tap or standpipe, tube well or borehole, protected dug well, protected spring and bottled water. Data on sanitation (proportion of households using improved toilet facilities in the district were obtained from the sanitation and environment department (or the planning office). The definition for 'improved toilet facilities' included: a flush toilet, a piped sewer system, a septic tank and a protected or covered or ventilated pit latrine. Data sources on transport and communication coverage in districts were obtained from the transportation and communication departments and/or the district planning office. The office of the district medical officer and the agricultural department were used as data sources on natural disasters occurring in the district.

7.2.6 Socioeconomic factors

Data on female literacy (as proxy for maternal education) were obtained from the Ministry of Education and Vocational Training. These data were limited to adult and non-formal education, primary education and secondary education. United Nations Children's Fund (UNICEF) defines literacy rate as "total number of literate persons in a given age group, expressed as a percentage of the total population in that age group".[221] According to UNICEF, adult literacy rate "measures literacy among

persons aged 15 years and older"[221] which matches with the conventional lower age for women of childbearing age. On the other hand, The World Bank defines adult female literacy rate as "the percentage of females age 15 and above who can, with understanding, read and write a short, simple statement on their everyday life".[222] According to either definition, the numerator for female adult literacy rate would be the total number of literate females aged 15 years or above. While the numbers obtained from the Ministry of Education and Vocational Training besides being useful and including all study districts, they did not include all females to satisfy the indicator requirement as there were no data for females with post-secondary education. The denominator for this indicator would also be the total number of females of age 15 years and above. There were no data available specific for the denominator for the study districts. As a proxy indicator of literacy in women of childbearing age, the study therefore used the numbers from the Ministry of Education and Vocational Training for the numerator and total population data available from the National Bureau of Statistics for the denominator.

7.2.7 Demographic factors

Contextual factors on demographic indicators (total fertility rate, infant and under-five mortality rates) were extracted from published reports from the National Bureau of Statistics.[k]

7.2.8 Transport and communication

Data on transport were obtained from the District's Transportation Department and/or the Planning Office and were based on the percentage (actual or likely) of villages/streets with accessibility to public (or own) transport services within 3 kilometres of district road's network. Data for transportation were of interest for linkage to reliable health facilities where antenatal care and obstetric care were provided. Data on communication (based on the actual or likely percentage of

---

k Total fertility rate, infant mortality rate, and under-five mortality rate data were obtained from the Tanzanian National Bureau of Statistics – the summaries of which appear in the Tanzania Demographic Health Survey, 2010 as aggregates of the administrative zones. The National Bureau of Statistics works with ICF Macro through its MEASURE DHS programme in conducting DHS surveys in Tanzania, please see also: www.nbs.go.tz and www.measuredhs.com

villages/streets covered by the available telecommunication services and networks in the district) were obtained from the District's Communication Department and/or the Planning Office. These type of data were needed for estimating a probable linkage between women (as users of FANC and EmOC services) and the services such as those provided through short messaging services (SMS) for communicating reproductive and maternal health information. Even though existence of such a linkage could not necessarily mean that women actually have access to reproductive and maternal health information, it was considered that the higher the coverage of mobile phone communication network (and potentially the more ownership of mobile phones among women) the more likely women could be associated with access to information.

7.2.9 Natural disasters

Data for the natural disasters (flood, famine, drought, and famine) were obtained from the District Medical Officer and/or the District's Agriculture Department. The main interest for collecting such data was to establish if a district low most affected with the natural disasters could be associated with weak implementation of FANC and EmOC programmes and/or low coverage of the services to women.

7.2.10 Completeness of data and analytical plan

Data for all proxy indicators were collected from all study districts except for Arusha Urban and Bagamoyo districts. Arusha Urban had data on female literacy levels and data on estimates on demographic indicators only. Data for other indicators from Arusha Urban could not be obtained during data collection period for reasons beyond the control of the study. Bagamoyo district had all data for contextual factors except for data on transport, communication, and natural disasters.

Data on transport for Bagamoyo districts were reported as not having been collected. The District Engineer for Bagamoyo advised to contact him after relevant surveys were completed later – outside the data collection period. Data on communication and natural disasters could not be obtained due to absence from office of the responsible personnel in the District Executive Director's office. In addition, Iringa Urban had missing data on the natural disasters of drought, floods, disease outbreaks, and famine. Following low missingness in data and using Acuna and

Rodriguez criteria,[155] all missing data for urban districts were imputed with urban-district arithmetic means and rural districts with rural-district arithmetic means. Data from questionnaires and data from other sources were centrally entered using Microsoft Excel 2007 and were transferred to Stata version 13 for analysis, where summary measures of means, medians, standard deviations (and 95% confidence intervals where applicable) – were generated.

Further analysis was conducted to compare differences between urban and rural districts in coverage of the indicators for the null hypothesis, H_0: "urban districts had similar levels of contextual factors as rural districts". Comparisons were done using a two-sample Wilcoxon rank-sum (Mann-Whitney) test. Unless otherwise stated, all P-values reported were those for the Wilcoxon rank-sum test comparing eight urban districts with 15 rural districts. The study made efforts to document the names of projects/initiatives that were in operation in each district. The list was however neither exhaustive of all related programmes/initiatives nor of the contextual factors. Programmes/initiatives documented were on maternal and newborn health services (other than FANC/EmOC), children health care services (immunization, micronutrients, nutrition, and so forth), malaria control programmes, HIV/AIDS services, and those on community health/social protection programmes.

7.3 Results

7.3.1 Total fertility, infant mortality, and <5 mortality rates

Results on demographic rates are shown in Table 7.1. Mean total fertility rates (TFR) in study districts in 2010 and 2011 were 5.18 (95% CI: 4.54 5.82) and 5.2 (95% CI: 4.46 5.78) respectively – slightly lower than the 2010 national estimates of TFR=5.4 reported in the Tanzania Demographic Survey.[161] Infant mortality rates (IMR) on the other hand, were 76.1 (65.9, 86.4) and 74.6 (64.6, 86.7) in 2010 and 2011 respectively. That is, for every one thousand live births that took place in study districts, about 76 babies in 2010 and 75 babies in 2011 died within their first year of life respectively. Study districts had much higher IMR than national estimates of 51 deaths per 1,000 live births. Under-five mortality rates (U5MR) were 120.9 (103.0, 138.8) and 118.2 (100.5, 135.8) for 2010 and 2011. U5MR were much higher in study districts compared to national estimates of 81 per 1,000 live births.

Table 7.1: Demographic factors on TFR, IMR, and U5MR for 2010 and 2011

No.	Study district	Total Fertility Rate		Infant Mortality Rate		Under 5 Mortality Rate	
		2010	2011	2010	2011	2010	2011
1	Arusha Urban	4.3	4.2	37.4	36.5	52.5	51
2	Ilala	3.2	3.1	60	58.6	93.1	90.6
3	Iringa Urban	3.97	3.84	102.8	101.1	166.4	163.2
4	Kinondoni	3.2	3.1	60	58.6	93.1	90.6
5	Mtwara Urban	4.83	4.81	109.8	108	181.2	177.9
6	Songea Urban	4.96	4.96	84.8	83.1	135.2	132.1
7	Tanga Urban	4.21	4.12	73.3	71.8	116	113.4
8	Temeke	3.2	3.1	60	58.6	93.1	90.6
	Urban Districts	4.0	3.9	73.5	72.0	116.3	113.7
9	Babati	4.3	4.2	37.4	36.5	52.5	51
10	Bagamoyo	4.99	4.95	87.4	85.8	139.7	136.7
11	Geita	5.7	5.6	76.9	75.3	120.9	118.2
12	Kahama	7.3	7.4	78	76.5	123.2	120.3
13	Kasulu	7.45	7.45	75.5	73.9	120	117.1
14	Kilosa	4.98	4.96	86.7	85	139.4	136.4
15	Kondoa	5.5	5.3	104.7	103	173	169.9
16	Mbozi	6.2	6.2	99.5	97.8	159.9	156.9
17	Moshi Rural	3.19	3.09	19.8	18.4	25	23.3
18	Muleba	7.37	7.38	96.7	95	157.8	154.7
19	Musoma Rural	7.03	7.04	94	92.2	153.1	149.9
20	Ruangwa	4.02	3.98	74.5	73	117.9	115.2
21	Singida Rural	5	4.84	61.7	60.3	95.6	93.1
22	Sumbawanga Rural	6.95	6.96	104	102.2	170.2	166.9
23	Uyui	7.27	7.28	66.1	64.7	102	99.4
	Rural Districts	5.8	5.8	77.5	76.0	123.3	120.6
	Overall	5.1	5	80.5	78.6	127.7	124.2
	Mean	5.2	5.1	76.1	74.6	120.9	118.2
	Standard Deviation	1.5	1.5	23.7	23.5	41.3	40.8

A comparison between urban districts and rural districts showed that the former had lower rates than the latter for both 2010 and 2011. However, only the lower rates in total fertility were significant with P-values of 0.003 (in 2010) and 0.005 (in 2011) – meaning that, women in urban populations will have fewer babies than their counterparts in rural districts. The leading districts with the lowest TFR rates in the two years of assessment were: Moshi Rural, Ilala, Kinondoni and Temeke, all with TFR=3.2 and all districts are from the Dar es Salaam region (the largest city in Tanzania) with the exception of Moshi Rural district. Districts with the highest rates in total fertility are Kasulu (7.45), Muleba (7.37) and Kahama (7.3), all of which are located in the western and Lake Victoria zones. Study districts with the lowest mortality rates for children (that is IMR and U5MR) in 2010 and 2011 were Moshi Rural, Arusha Urban and Babati. Districts with the highest mortality rates for children were Mtwara Urban, Kondoa, Sumbawanga Rural, and Iringa Urban.

7.3.2 Female literacy levels

Results on female literacy levels are shown in Table 7.2. Overall, the number of females with adult and non-formal education was reported to be higher across the districts than the number of those with primary and secondary education. The female literacy proportion, defined here as the sum of the number of females with adult and non-formal education, primary education and secondary education divided by the annual population, was reported to be higher in 2010 (mean=22.2%, 95% CI: 21.1%–23.3%) than in 2011 (21.4%; 20%–22.8%). The female literacy proportion for rural districts appeared slightly higher than for urban districts in the year 2010, with a mean percent of 22.4% (for rural districts) and 21.8% (for urban districts). Female literacy proportions for 2011 were similar for both rural and urban districts. The difference in female literacy proportion observed in 2010 between rural and urban districts was however not significant (P-value=0.846). Geita, Arusha Urban, and Musoma Rural had the highest female literacy levels, while to Ruangwa, Kinondoni, and Mtwara Urban that had the lowest female literacy levels in 2010. In 2011, Arusha Urban, Geita, and Temeke had the highest female literacy levels, and Ruangwa, Musoma Rural, and Songea Urban had the lowest.

7.3.3 Water and sanitation

Table 7.3 shows the percentage of households in study districts that used drinking water from improved water sources and those using improved toilet facilities. The overall coverage of the two indicators was higher in 2011 than it was in 2010. The percentage of households using drinking water from improved water sources rose from a mean of 58.7% (95%CI: 50.3%–67%) in 2010 to 61.9% (53.2%–70.6%) in 2011. Use of improved toilet facilities also rose from a mean of 50.1% (40.0%–60.1%) in 2010 to 55.5% (46.3%–64.6%) in 2011. Urban districts had higher coverage of water and sanitation compared to rural districts. About three quarters (74.6%) of households in urban districts in 2010 used water from improved water sources, compared to just above half (51.3%) of households in rural districts. In 2011, the proportion of households using water from improved water sources in urban districts was 78.4% compared to 54.2% in rural districts.

Table 7.2: Socioeconomic factors on proportion of literate females in study districts for 2010 and 2011

No	Study district	Number populations in study districts		Number of females with adult and non-formal education		Number of females with primary education		Number of females with secondary education		% female literacy to district pop	
		2010	2011	2010	2011	2010	2011	2010	2011	2010	2011
1	Arusha Urban	362,484	372,719	35,992	37,068	30,847	35,992	26,948	30,129	25.90%	27.7%
2	Ilala	722,027	730,108	76,751	78,445	74,438	76,751	21,097	25,069	23.90%	24.7%
3	Iringa Urban	152,649	158,592	13,206	13,464	11,659	13,206	6,727	6,628	20.70%	21.0%
4	Kinondoni	1,237,145	1,250,230	94,692	94,194	93,607	94,692	27,847	31,338	17.50%	17.6%
5	Mtwara Urban	122,588	126,923	8,768	8,700	8,733	8,768	3,612	3,431	17.20%	16.5%
6	Songea Urban	175,658	182,156	19,547	20,821	17,047	19,547	5,252	5,672	23.80%	25.3%
e	Tanga Urban	298,881	305,713	26,220	26,568	30,786	26,220	8,959	8,520	22.10%	20.1%
8	Temeke	870,999	880,022	88,845	88,895	88,506	88,845	24,682	26,003	23.20%	23.2%
9	Babati	407,780	423,380	38,416	38,176	39,157	38,416	8,748	10,223	21.20%	20.5%
10	Bagamoyo	277,673	283,780	27,995	27,528	27,367	27,995	7,180	8,023	22.50%	22.4%
11	Geita	856,074	879,836	112,775	116,368	108,203	112,775	9,058	12,747	26.90%	27.5%
12	Kahama	815,177	848,728	77,847	75,755	77,639	77,847	7,067	7,427	19.90%	19.0%
13	Kasulu	631,318	653,858	64,994	60,272	63,976	64,994	7,360	7,245	21.60%	20.3%
14	Kilosa	587,967	600,428	54,014	54,056	54,972	54,014	6,185	7,909	19.60%	19.3%
15	Kondoa	499,407	508,353	54,063	55,891	55,652	54,063	8,356	7,048	23.60%	23.0%
16	Mbozi	662,287	681,969	76,721	72,460	76,118	76,721	12,574	12,665	25.00%	23.7%
17	Moshi Rural	462,085	469,593	39,183	38,343	40,714	39,183	20,418	18,576	21.70%	20.5%
18	Muleba	475,742	492,404	49,857	51,767	51,544	49,857	5,918	7,273	22.60%	22.1%
19	Musoma Rural	419,962	669,365	49,231	48,706	49,014	49,231	7,323	6,921	25.10%	15.7%
20	Ruangwa	146,990	149,634	12,066	11,592	13,102	12,066	632	1,380	17.60%	16.7%
21	Singida Rural	486,901	497,562	55,434	46,917	47,835	55,434	8,342	6,313	22.90%	21.8%
22	Sumbawanga R	482,987	499,587	52,690	49,326	56,431	52,690	4,625	4,336	23.60%	21.3%
23e	Uyui	352,560	362,619	40,199	36,122	37,943	40,199	1,756	2,190	22.70%	21.7%
	Urban Districts		Mean (Median)	45,503 (31,106)	46,019 (31,818)	44,453 (30,817)	45,503 (31,106)	15,641 (15,028)	17,099 (16,795)	21.8% (22.7%)	21% (22.1%)
	Rural Districts		Mean (Median)	53,699 (52,690)	52,219 (49,326)	53,311 (51,544)	53,699 (52,090)	7,703 (7,323)	8,018 (7,273)	22.4% (22.6%)	21% (21.3%)
	Overall		Mean (95% CI) (Median)	50,848 (38,847 62,849)	50,062 (38015 62,110) (48,706)	50,230 (38,447 62,013)	50,848 (38,847 62,849) (49,857)	10,464 (7,034 13,894)	11,177 (100.5, 135.8) (7,427)	22.2% (21.1%, 23.3%)	21.4% (20%, 22.8%)
	Test for H_o; urban =rural		P-value	0.366	0.439	0.333	0.366	0.156	0.197	0.846	0.519

Table 7.3: Water and sanitation factors: Proportion of households using drinking water from improved water sources and improved toilet facilities in 2010 and 2011

No.	Study district	Residence	% hholds using drinking water from improved water sources		% hholds using improved toilet facilities	
			2010	2011	2010	2011
1	Arusha Urban	Urban	75	78	69	73
2	Ilala	Urban	85	90	80	85
3	Iringa Urban	Urban	93	95	75	85
4	Kinondoni	Urban	70	75	55	55
5	Mtwara Urban	Urban	52	58	48	50
6	Songea Urban	Urban	71	74	94	98
7	Tanga Urban	Urban	86	87	80	82
8	Temeke	Urban	65	70	50	55
9	Babati	Rural	59	61	25	35
10	Bagamoyo	Rural	68	68	39	43
11	Geita	Rural	55	59	44	46
12	Kahama	Rural	58	64	44	49
13	Kasulu	Rural	68	80	70	62
14	Kilosa	Rural	75	79	78	80
15	Kondoa	Rural	32	34	50	62
16	Mbozi	Rural	38	39	15	48
17	Moshi Rural	Rural	73	74	22	40
18	Muleba	Rural	60	65	52	52
19	Musoma Rural	Rural	39	43	12	13
20	Ruangwa	Rural	36	36	63	67
21	Singida Rural	Rural	32	32	40	40
22	Sumbawanga	Rural	48	49	41	43
23	Uyui	Rural	28	30	25	30
Urban Districts		Mean (Median)	74.6% (71%)	78.4% (75%)	68.8% (75%)	72.96% (82%)
Rural Districts		Mean (Median)	51.3% (55%)	54.2% (59%)	41.3% (41%)	47.4% (46%)
Overall		Mean (95% CI) (Median)	58.7% (50.3% 67%) (59.5%)	61.9% (53.2% 70.6%) (64.5%)	50.1% (40% 60.1%) (49%)	55.5% (46.3% 64.6%) (51%)
Test for urban=rural		P-value	0.01	0.009	0.006	0.005

The percentage of households using improved toilets in urban districts in 2010 was about 69% compared to about 41% of households in rural districts. Similarly, in 2011, about 73% of households in urban districts used improved toilets, compared to 55.5% in rural districts. The evidence of the differences in household use of water from improved water sources and use of improved toilets between urban and rural districts was strong, with all P-values less than or equal to 0.01. Iringa Urban, Tanga Urban, Ilala, and Songea Urban (all urban districts) had an overall highest coverage in water and sanitation for the years 2010 and 2011. Moshi Rural was a special district in that it had the highest coverage of households using water from improved

water sources (98% in 2011) but had less than a quarter and only 40% of its households using improved toilet facilities in 2010 and 2011 respectively.

7.3.4 Transport and communication

Results on transport and communication indicators are shown in Table 7.4. The percentage of villages/streets with access to public transport services within three kilometres of the district's road network increased from 65% (95%CI: 55%–74%) in 2010 to 70% (62%–78%) in 2011. However, not all roads within a district's road network are passable throughout the year. During the rainy seasons of 2010 and 2011 respectively, 30% and 26% of villages/streets had their roads not passable. On the other hand, telecommunication services coverage in study districts' villages/streets went up from 59% (44%–75%) in 2010 to 63% (48%–78%) in 2011. Telecommunication services include use of mobile and landline phones, and are provided by the Tanzania Telecommunication Company Limited – the main provider of landline phone services. Main providers of mobile phone services are Vodafone/Vodacom, Airtel, Tigo, and Zantel.

There were notable differences between urban and rural districts in coverage of transport and communication services, although evidence on such differences was very weak, P-values >0.05. Based on access to public transport within three kilometres of districts' road networks, urban districts had higher coverage than rural districts: 67% versus 63% in 2010, and 73% versus 68% in 2011. Likewise, telecommunication service coverage in villages/streets was better in urban districts compared to rural districts: 73% versus 52% in 2010, and 77% versus 56% in 2011. The urban districts of Songea, Tanga, and Iringa had, on average, the highest coverage of their streets with transport services. Districts with the highest coverage in communication services in villages/streets were those of Ilala, Songea Urban, and Moshi Rural. Of particular interest, even though Kahama district is a rural district, it also ranked high in communication service coverage of its villages. The high coverage could potentially be linked to the presence of goldmines that have over the last decade attracted international and local companies to operate in the district, thereby increasing mobile phone companies' presence.

Table 7.4: Infrastructural factors on access within districts to transport and communication networks for 2010 and 2011

No	Study district	Residence	% villages/streets with access to public transport within 3km		% of villages/streets with impassable roads during rainy season		% of villages/streets covered by available telecomm services	
			2010	2011	2010	2011	2010	2011
1	Arusha Urban	Urban	67	73	20	15	73	77
2	Ilala	Urban	75	80	20	18	100	100
3	Iringa Urban	Urban	80	85			50	60
4	Kinondoni	Urban	60	65	40	30	80	80
5	Mtwara Urban	Urban	24	32	3	5	24	32
6	Songea Urban	Urban	95	97	0	0	100	100
7	Tanga Urban	Urban	84	95	16	5	86	87
8	Temeke	Urban	50	60	40	30	70	80
9	Babati	Rural	15	40	44	45	38	48
10	Bagamoyo	Rural	63	68	35	32	52	56
11	Geita	Rural	32	35	13	12	54	64
12	Kahama	Rural	76	81	20	16	92	94
13	Kasulu	Rural	89	72	63	72	0	0
14	Kilosa	Rural	56	67	21	8	77	78
15	Kondoa	Rural	70	80	30	20	90	90
16	Mbozi	Rural	61	63	44	40	10	13
17	Moshi Rural	Rural	74	81	12	10	94	96
18	Muleba	Rural	55	60	56	58	50	55
19	Musoma Rural	Rural	75	78	25	22	73	85
20	Ruangwa	Rural	80	87	20	13	0	0
21	Singida Rural	Rural	65	65	35	35	55	55
22	Sumbawanga R	Rural	69	72	26	22	46	48
23	Uyui	Rural	70	70	75	75	0	0
Urban Districts		Mean (Median)	67% (75%)	73% (80%)	20% (18%)	15% (12%)	73% (80%)	77% (80%)
Rural Districts		Mean (Median)	63% (70%)	68% (71%)	35% (28%)	32% (22%)	52% (54%)	56% (55%)
Overall		Mean (95% CI) (Median)	65% (55% 74%) (70%)	70% (62% 78%) (72%)	30% (21% 39%) (26%)	27% (17% 37%) (21%)	59% (44% 75%) (63%)	63% (48% 78%) (71%)
Ho; urban =rural		P-value	0.55	0.478	0.116	0.083	0.177	0.142

7.3.5 Floods, drought, outbreaks, and famine

Table 7.5 shows results on the extent districts were affected by natural disasters (floods, drought, disease outbreaks, and famine) in 2010 and 2011. About a fifth (19%, 95%CI: 7%–31%) and 22% (10%–33%) of villages/streets in study districts were affected by drought in 2010 and 2011 respectively. Floods in study districts affected 9% (2%–16%) of villages/streets in 2010 and 13% (3%–24%) of villages/streets in 2011. About a tenth of villages/streets in study districts were affected by disease outbreaks (9%, 2%–16%, in 2010; and 10%, 3%–18% in 2011).

Table 7.5: Percentage of households affected by natural disasters in 2010 and 2011

No	Study district	% of district's households affected by:							
		Drought		Floods		Disease Outbreaks		Famine	
		2010	2011	2010	2011	2010	2011	2010	2011
1	Arusha Urban	12	12	8	26	6	11	1	2
2	Ilala	0	0	5	5	2	1	0	0
3	Iringa Urban	12	12	8	26	6	11	1	2
4	Kinondoni	0	0	0	50	0	30	0	0
5	Mtwara Urban	61	64	31	43	2	4	1	2
6	Songea Urban	0	0	0	0	0	0	0	0
7	Tanga Urban	0	0	0	0	0	0	0	0
8	Temeke	10	10	13	30	30	30	5	7
9	Babati	6	6	0	0	0	0	25	25
10	Bagamoyo	22	26	10	8	11	10	22	24
11	Geita	5	2	1	1	0	0	5	2
12	Kahama	36	45	1	0	0	1	14	20
13	Kasulu	0	0	0	0	0	0	0	0
14	Kilosa	32	34	46	10	12	7	35	45
15	Kondoa	46	60	15	20	38	14	35	40
16	Mbozi	5	7	2	1	5	8	5	7
17	Moshi Rural	1	3	0	0	0	0	6	4
18	Muleba	70	75	50	58	50	60	70	75
19	Musoma Rural	45	48	18	15	13	16	75	78
20	Ruangwa	0	0	0	0	0	0	0	0
21	Singida Rural	65	55	0	0	3	4	35	35
22	Sumbawanga R	0	30	0	1	28	31	0	0
23	Uyui	0	0	0	0	0	0	0	0
	Urban Districts Mean	12%	12%	8	26%	6%	11%	1%	2%
	(Median)	(0%)	(0%)	(3%)	(18%)	(1%)	(3%)	(0%)	(0%)
	Rural Districts Mean	22%	26%	10%	8%	11%	10%	22%	24%
	(Median)	(6%)	(19%)	(1%)	(1%)	(2%)	(3%)	(10%)	(14%)
	Overall Mean	19%	22%	9%	13%	9%	10%	16%	17%
	(95% CI)	(7% 31%)	(10% 33%)	(2% 16%)	(3% 24%)	(2% 16%)	(3% 18%)	(5% 26%)	(5% 26%)
	(Median)	(5%)	(7%)	(1%)	(1%)	(1%)	(3%)	(5%)	(3%)
	Ho; urban =rural (P-value)	0.233	0.178	0.965	0.194	0.623	0.865	0.033	0.033

Figure 7.1: Dar es Salaam's infrastructural and residential devastation by floods (Taken on 22nd Dec 2011, courtesy of Deutsche Welle Swahili Service)[223]

Likewise, in 2010, famine affected study districts an average of 16% (5%–26%) of villages and streets, and 17% (5%–26%) in 2011. Urban districts were least affected by famine compared to the other natural disasters (only 1% in 2010 and 2% in 2011). Floods affected 12% of streets in urban districts in both 2010 and 2011. In particular, some parts of Kinondoni and Ilala districts in the main city of Dar es Salaam were most affected by the December 2011 floods. Mainstream media reported at least 23 people died and thousands of households were left homeless[224 225]. Figure 7.1 shows pictures of the effect of floods after devastating households and infrastructure on 22nd December 2011. No disease outbreaks were reported as following the floods.

7.3.6 Other non-FANC/EmOC projects/initiatives

Table 7.6 shows a list of projects and initiatives that were identified operating in study districts in 2010 and 2011. It also shows the purposes of those projects and initiatives, including the main sources of funding and a general overview of coverage.

7.4 Discussion

This chapter presented results of the levels of contextual factors that could be associated with the implementation strength and coverage of FANC and EmOC programmes in the 23 study districts of Mainland Tanzania. There is a large contextual variation across study districts whose relevance should be considered in explaining the extent of the differences in outcomes. Such levels of contextual factors should be accounted for in statistical analyses as factors that can potentially confound the association between the programme implementation strength and programme results. Even if there would be no statistical difference between urban and rural districts for example, it is possible that contextual factors are likely to affect either the implementation of programme activities or programme outcomes or both and should thus be analysed and reported. In an attempt to explore their role as potential confounders in the association between implementation strength

Table 7.6: Public health system factors based on other projects and initiatives operating in study districts in 2010 and 2011

Types and purposes of projects/initiatives that were in operation in study districts in 2010 and 2011 that could potentially affect FANC and EmOC outcomes:

- Maternal and newborn health services (other than FANC/EmOC):

Purposes: Family planning, provision of youth-friendly services, development of skills for health workers, provision of supplies and drugs to health facilities, payment for performance aimed at improving provision of maternal health services, prevention of mother to child transmission of HIV (PMTCT), reduction of maternal and child deaths, advocacy for male involvement to attend ANC clinics, outreach provision of vaccines and supplies, renovation of health facilities, and many others.

- Child health care services: (immunization, micronutrients, nutrition, etc.)

Purposes: providing nutrition for children in primary schools, reducing infections to infants, early infant diagnosis, food fortification, child protection against illnesses, reduction of malnutrition for mother and child, improving newborn survival, nutrition for women and newborns, distribution of nutritional food to children.

- Malaria control programmes:

Purposes: Larval control of malaria vector mosquitoes in their breeding sites, provision of subsidized vouchers for purchasing ITN bed nets for pregnant women and newborns, distribution of mosquito bed nets, promotion of proper hang up of bed nets and conducting research on the quality of the insecticide for spraying.

- HIV/AIDS services: (most widespread in all districts and with more coverage in both health facilities and communities)

Purposes: Voluntary and provider-initiated counselling and testing of HIV/AIDS, provision of care and treatment and home-based care for people living with HIV/AIDS, Prevention from Mother To Child Transmission of HIV, friendly services for youth reproductive health, service for TB/AIDS, reduction of HIV infection through male circumcision, training of health care providers in PMTCT, supply of condoms and other equipment/supplies, providing psychosocial support, community mobilization for TB care, cancer screening for HIV/AIDS patients, prevention and management of STI/RTI, early infant diagnosis community sensitisation on HIV/AIDS, training of health workers and supply of HIV/AIDS drugs.

- Community health/Social protection programmes

Purposes: Development and delivery of services to the most vulnerable children, sensitization on care for water sources and the environment, community health fund for health access, caring for the elderly, building residential houses for and providing loans and educational support to most vulnerable children, supporting disabled children with bicycles and educational materials, supporting children and adults in acquiring birth certificates, early childhood care and development, enhancing access to justice by the poor and the most vulnerable, provision of clean and safe drinking water, health promotion and strengthening of the health system, home-based care for the disabled, building capacity for children for essential primary services.

Funding received from:

- Local initiatives/programmes:

National Malaria and AIDS Control Programmes (Ministry of Health and Social Welfare), Comprehensive Community-Based Rehabilitation in Tanzania (CCBRT), Uzazi na Malezi Bora Tanzania (UMATI), African Medical and Research Foundation (AMREF), Wanawake na Maendeleo (WAMA), Hapa Project, "BMC"/Ministry of Health and Social Welfare, Wazazi na Mwana, Tanzania Food and Nutrition Centre, Tanzania Red Cross Society, The Rotary Club of Tanga, Geita Gold Mines (Anglo Ashanti Gold), Angaza Project, Mkombozi Project, "BADO" project and Moyo Mmoja Trust.

- USAID- or CDC-funded projects:

Management and Development for Health (MDH), Engender Health, Public Services International (PSI), Marie Stoppes, Jhpiego Corporation, Elizabeth Glaser Paediatric AIDS Foundation (EGPAF), Aids Relief, International Center for AIDS Care and Treatment (ICAP), Family Health International, Mennonite Economic Development Associates (MEDA), Helen Keller/MOHSW, Joining Hands Initiative (JHI), Urban Research and Consulting (URC), Johns Hopkins University's Center for Communication Programs (JHU-CCP), USA President's Malaria Initiative (PMI), USA President's Emergency Plan for AIDS Relief (PEPFAR), Pathfinder, IntraHealth, Water Reed, Management Sciences for Health (MSH), The International Training and Education Center for Health (*I-TECH*), PharmAccess, PATH Global Health Organization, Campaign for Female Education (CAMFED) and PACT World.

- Other international projects/initiatives:

Bill and Melinda Gates Foundation, The Global Fund for TB, HIV/AIDS and Malaria, Clinton Health A Initiative (CHAI), World Vision, CARE International, Africare, Baylor Paediatric Clinic, Save the Children, Japan International Cooperation Agency (JICA), London School of Hygiene and Tropical Medicine (LSHTM), Swedish International Development Cooperation Agency (SIDA), Canadian International Development Agency (CIDA) and Plan International.

- United Nations programmes:

United Nations Population Fund (UNFPA), United Nations Children's Fund (UNICEF), United Nations World Food Programme (WFP).

Serviced areas/coverage:

Community-based projects/initiatives covered as little as one ward in a district to nearly all wards. Health facility-based projects/initiatives provided services in as few health facilities as two dispensaries to as many as all health facilities in the district. For example the Cancer Screening Project in Iringa Urban district operated in one health facility in 2010 and 2011 while the ITN Voucher project in Tanzania for provision of vouchers for use in purchasing subsidized insecticide-treated bed nets for pregnant women and newborns operated in all 118 health facilities of Sumbawanga Rural districts in 2010 and 2011, as did the PARMVEC project in Muleba district, which conducted research on the quality of the insecticide for spraying in all villages of the district.

components and programme results, univariable regression analyses were conducted in a separate chapter (Chapter 8 on illustrative dose-response) for illustration only.[1]

Overall, female literacy levels in study districts were similar in 2010 and 2011, with 22.2% and 21.4% respectively. A number of studies have shown that mother's education is associated with increased use of maternal health services, which is linked to reduction in maternal deaths.[226] [227] [228] For example, a study in 2010 assessed the significance of 'national female literacy' in use of maternal health service in sub-Saharan countries and found that higher national female literacy levels were associated with decreased 'income-related inequalities' in use of maternal health services.[229] Obtaining data on female literacy levels was relatively difficult and was only possible after the author visited a colleague at the Ministry of Education and Vocational Training who by chance directed him to the person with data for all study districts. In reflection, this points out to a more general problem of public institutions improving the availability of information on public goods and services.

Study districts showed some signs of improvements in water and sanitation in 2011 compared to 2010. The percentage of households using drinking water from improved water sources rose from 59% in 2010 to 62% in 2011. Likewise, use of improved toilet facilities in study districts also rose from 50% in 2010 to 56% in 2011. There was a slight increase in coverage of telecommunication services in villages/streets within study districts in 2010 and 2011. Private mobile phone companies are scaling up their phone coverage networks to reach more customers in both urban and rural areas. Tanzania has pioneered and is currently implementing a significant number of health interventions using mobile phone services, aimed at improving access to health care and reaching more populations.

While there are more than 30 organizations in Tanzania (most being public-private partnerships) using such service-based approaches,[230] two examples are worth

---

[1] Note that the contextual factors presented in this chapter were single point observations without change over time data. To be able to estimate their confounding effect in a multiple regression model, it is necessary that both programme coverage (response-side) and contextual factors include change over time data that I was not able to collect due to time and resource limitations.

naming: the SMS for Life project[231] and the Registration, Insolvency and Trusteeship Agency's (RITA) birth registration system.[232] SMS for Life started as a small project to pilot stock management of anti-malaria drugs in 129 health facilities in three districts of rural Tanzania through use of 'simple and widely available SMS technology'.[231] Following its success, the SMS for Life project is currently monitoring anti-malaria drug stock levels in most health facilities in the country. RITA is the government's agency for vital registration and, is currently collaborating with Tigo – one of the country's mobile phone operators to scale-up registration of births in Mainland Tanzania. The registration of birth involves use of a mobile phone application to send some information about the birth to RITA's central database.

In addition to Tanzania's increased mobile phone use in providing solutions in various fields, an evaluation of a recent study in Zanzibar reported that mobile phones usage was also linked with increased antenatal care visits used among pregnant women.[233] Further, the "Wazazi Nipendeni" (meaning "Parents Love Me") countrywide initiative is a multi-media campaign led by the Tanzanian Ministry of Health and Social Welfare in partnership with several maternal health stakeholders with the aim to promote healthy pregnancy and safe motherhood through use of mobile phone text messages to receive the information.[234] Elsewhere, use of mobile phone services has been used in disease surveillance, sharing SMS messages for reproductive health and transferring money (mobile money) and in telemedicine through texting reminder messages for clients to attend for health services at a health facility.[235] Improved communications in study districts therefore has the potential to improving access to FANC and EmOC programmes

The percentage of villages/streets in study districts with access to public transport services within three kilometres of district's road network increased from 65% in 2010 to 70% in 2011. However, about 30% of study districts' villages/streets had their roads impassable during rainy seasons in 2010, and about a quarter in 2011. Transport (in terms of both distance, travel time and cost) to health facilities has been linked as a contributing factor in Thaddeus and Maine's first two delays: delayed decision to seek care and delayed arrival at a health facility.[130] A woman with an emergency obstetric complication may find difficulty in getting to a facility where she can get the needed care. Distance to health facilities has been found to decrease use of health facilities, including women's attendance at antenatal care clinics.[26 236] Compared to most countries south of the Sahara, Tanzania is fortunate

to have a network of health facilities, with about three-quarters of its population residing within 5 kilometres. [237] [238] While this is the case, transport in rural Tanzania is still a challenge to many populations and users of maternal health care. Therefore, evaluating study districts on access to public transport services within district's road network is essential in an attempt to establishing the likelihood of villages/streets vicinity to health facilities.

About a fifth of villages/streets in study districts were affected by drought. On average, floods in study districts affected 9% of villages/streets in 2010 and 13% in 2011. About a tenth of villages/streets in study districts were affected by disease outbreaks. Likewise, famine affected about one sixth of villages and streets in 2010 and 2011. Natural disasters have been associated with displacement of populations rendering them to seek health care and humanitarian relief.[239] Natural disasters in low-resource countries have a bigger effect than those in wealthier countries due to the lack of or weakness of systems for disaster preparedness.[240] Effects of natural disasters are mostly grouped in three categories[241] – health, social and economic – and they tend to affect mostly women and children in the afflicted areas.[242] In 2009, a collaborative commission between The Lancet and University College London, UK[243] was set up to evaluate climate change and its effect on health over the coming decades. The commission identified climatic effect on health as lack of food and safe drinking water, poor sanitation, population migration, changing disease pattern and morbidity, more frequent extreme weather events, and lack of shelter.

It was challenging to convert the number and types of documented programmes and initiatives (other than FANC and EmOC) that were in operation within each district into one or more quantitative measure for use in comparing study districts. Nearly all study districts had some of these projects running. Due to limitation of not having 'change-over-time' data, it was practically not possible to estimate the impact of the contextual factors on FANC and EmOC programmes. However, besides not having been quantitatively accounted for, each of the programmes and initiatives locally operating in study districts had their contribution in facilitating the implementation of the two programmes of FANC and EmOC. For example, for the FANC programme, the indicator on HIV testing had most of its coverage actually reached through the Prevention of Mother to Child Transmission of HIV/ADIS (PMTCT) vertical programme that has enjoyed nearly universal coverage in Tanzania, which receives

most of its financial support from PEPFAR (United States President's Emergency Program for AIDS Relief).

7.5 Limitations

Design issues should be taken into consideration when using these findings. Use of district officials in various departments of the District Executive Director could have attracted some reporting bias as they responded to survey questions. For example, it is possible that district engineers could have introduced some errors in estimating the telecommunications and transportation coverage within their districts. Even though most data were documented, some data may have been provided without a basis in evidence, especially in estimating the effect of natural disasters.

Assessment of the contextual factors was based on percentages of households and/or villages and streets. Percentages were only used to describe the status of the factors in study districts and no other measures of associations were used. It is therefore possible that including other measures of assessment could have highlighted other findings other than those presented here. As was mentioned in Section 7.4 above, this study was further limited by not having estimated the impact (such as a composite indicator) of other programmes operating in study districts on FANC and EmOC programmes. It is therefore possible that study districts would have had different levels of implementation strength (Chapter 4) and programme coverage (Chapter 6) if these other programmes operating in districts would have been estimated.

7.6 Conclusion

This chapter has summarized the process that was used to identify the types of contextual factors related to FANC and EmOC programmes in study districts as well as how the factors were measured and analysed. While it was relatively easy to obtain some types of the contextual factors (such as total fertility rate, infant and under-five mortality rates, water and sanitation, and transport and communication, it was relatively difficult to obtain data on natural disasters (floods, drought, famine, and disease outbreaks) as well as female literacy levels for all districts. For the natural disasters data, it is important that districts should continually document

their levels and the effects caused by these for purposes of planning and other use. In addition, data on other projects and initiatives operating in districts was streamlined well at the district medical officer's office – reflecting a need for improving documentation and sharing of such information. Overall, collection of data for the contextual factors was variable, being relatively easy for some issues and challenging for others. Further investigation is still needed in finding reliable ways of estimating the composite index of the contextual factors for inclusion in regression analyses. As there are numerous such programmes and initiatives operating in districts, understanding the magnitude of their effect in the implementation and results of FANC and EmOC programmes is of great importance. Furthermore, to increase the credibility of results, qualitative information is also required on the contextual factors from users of the services, which is more likely to shed more light underlying issues facing users of FANC and EmOC programmes as opposed to quantitative information reported in this study.

Chapter 8: Illustrative dose-response analysis

8.1 Introduction

In epidemiology, Bradford Hill outlined nine conditions that are likely to be observed if there is a causal relationship between a cause (exposure) and effect (outcome).[244] In two of the nine conditions, Hill suggested that a causal relationship is more likely if the association between a cause and the effect is strong and if more of the cause (dose) results in more of the effect (response). This is demonstrated in an example of an American study that investigated the association between body mass index (BMI) before pregnancy and pre-defined birth defects.[245] The study found that, compared to women with normal weight (BMI=18.5–24.9), where the prevalence of birth defects among live born babies was 3.7%, the prevalence of birth defects in live births increased to 12% in obese women (BMI=30.0–34.9), 26% in obese women (BMI=35.0–39.9), and 37% in very obese women (BMI≥40.0). While Hill's criteria have been widely applied in establishing causal relationships between causes and effects, Rothman and Greenland pointed out that Hill's criteria have a limited application in evaluating public health interventions because of the complex nature of causal factors on the outcome being evaluated.[246]

The main objective of this chapter was to investigate the nature of the associations between programme results (coverage – as responses) and the implementation of FANC and EmOC programme components (as doses), using results from previous chapters. This analysis also looked at the associations between programme results and the contextual factors in univariable regression analysis. This analysis however used data that were based on a single point in time, thus assuming the programme implementation and confounders are constant over time, and ignoring the changes in the programme response over the different time intervals. Most of the indicators of the implementation strength, programme results, and the contextual factors had data for year 2011 only, with relatively few indicators having data for both years (2010 and 2011). Due to lack of data from two (or more) time points, results presented in this chapter are therefore illustrative of the proposed method – which when fully

implemented should have data from continuous monitoring of programme activities throughout programme lifetime. In order to carry out a dose-response analysis in a programme, the outcome (response) would be the programme results at several different time-points, while the dose would be the implementation strength, again measured at several time-points. The potential confounders would also be contextual factors that change over time and for which the change over time may be associated with the response.

8.2 Analytical approach

Data analysis was performed in two stages: stage one involved conducting univariable regression analyses and step two involved graphical presentation of the dose-response results. In stage 1, univariable regressions were used to explore the relationship between individual programme implementation strength variables and the overall coverage scores of FANC and EmOC programmes. For FANC programme, the overall coverage score on the response side was 'FANCcov' (appearing as 'district score' in Table 6.1) which was the mean score of the three FANC programme aggregate indicators: Antenatal care coverage, Tetanus toxoid coverage, and HIV screening coverage. Likewise, for EmOC programme, the overall coverage score on the response side was 'EmOCcov' (appearing as 'district score' in Table 6.2) which was the mean score of the three aggregate indicators of EmOC availability: Average Signal Function Score, Institutional delivery rate, and Caesarean section rate.

Programme implementation strength variables were the six programme components of human resources, service delivery, drugs and supplies, health financing, health information system and leadership and governance whose scores for each district appear in Table 5.2 for FANC programme (giving the FANCscore) and Table 5.3 for EmOC programme (giving the EmOCscore). Univariable regression analyses also included contextual factors of total fertility rate, infant mortality rate, under five mortality rate, female literacy levels, water, sanitation, transport, communication, floods, drought, disease outbreaks, and famine, which are considered as potential confounders of programme coverage. Scores for the contextual factors are shown in Tables 7.1 through 7.5 with averages taken for the years 2010 and 2011. In stage 2, dose-response graphs were generated, with one data-point for each district, using the composite score of the programme implementation strength components (FANCscore and EmOCscore) as doses in the x-axis, against the overall coverage

scores of programmes (FANCcov and EmOCcov) as responses in the y-axis. One graph was generated for the FANC programme, and one graph for the EmOC programme. For illustration purposes, multivariable analysis were conducted and included predictor variables (programme components and contextual factors) that showed significance associations at p = 0.05 in the univariable analysis stage. For FANC, predictor variables were: human resources, FANC service delivery, health financing, female literacy, and sanitation; and for EmOC, predictor variables were: human resources, EmOC service delivery, EmOC drugs and supplies, health financing, total fertility rate, water, sanitation, and famine. All regression analyses and generation of graphs were conducted in Stata version 13.

8.3 Results

8.3.1 Univariable analysis

Tables 8.1 and 8.2 present the results from univariable analysis of the associations between overall programme coverage scores of FANC and EmOC and their corresponding programme implementation strength scores and the contextual factors. Three of the six programme components were significantly associated with the overall FANC programme coverage and included the following components: human resources (P=0.008), FANC service delivery (P=0.009), and health financing (P=0.042). Only two of all contextual factors were significantly associated with the overall FANC programme coverage (P<0.05). These were: female literacy levels (adult, primary and secondary education, P=0.004), and sanitation (households using improved toilet facilities, P=0.029).

With the exception of the health information system and leadership and governance components, all other programme components were significantly associated with the overall coverage of EmOC programme (human resources, P=0.003; service delivery, P=<0.001; drugs and supplies, p0.006; and health financing, P=0.001). In addition, three contextual factors were significantly associated with the overall coverage of EmOC programme and included total fertility rate (P=0.03), water (households using drinking water from improved water sources, P<0.001), and sanitation (households using improved toilet facilities, P<0.001). There was weak evidence that famine was associated with the overall EmOC programme coverage (with borderline p = 0.055)

and was therefore included in the second stage of multiple regression other programmes operating in study districts.

Table 8.1: Univariable analysis for associations between the Overall FANC Coverage and the FANC Components & Contextual Variables in 23 districts

Predictor variables	Slope (95% CI)	P-value
Programme implementation strength components	FANCcov*	
Human resources	0.614 (0.176 – 1.051)	0.008
Service Delivery	0.83 (0.234 – 1.427)	0.009
Drugs and Supplies	0.269 (-0.1762 – 0.713)	0.223
Health Financing	0.484 (0.02 – 0.948)	0.042
Health Information System	0.397 (-0.703 – 1.497)	0.461
Leadership and Governance	0.447 (-0.705 – 0.795)	0.903
Overall FANC Implementation Strength Score Model: **FANCcov = 0.358 + 1.007*FANCscore$^{\beta}$**	1.007 (0.231 – 1.783)	0.013
Contextual factors		
Total Fertility Rate	-0.024 (-0.101 – 0.0526)	0.517
Infant Mortality Rate	-0.001 (0.006 – 0.003)	0.547
Under five Mortality Rate	-0.001 (-0.004 – 0.002)	0.516
Female literacy levels (adult, primary and secondary education)	5.501 (1.989 – 9.014)	0.004
Water (households using drinking water from improved water sources)	0.425 (-0.159 – 1.006)	0.143
Sanitation (households using improved toilet facilities)	0.551 (0.061 – 1.042)	0.029
Transport (accessibility to public transport services within 3 kilometres; pass-ability of roads in rainy season)	0.254 (-0.679 – 1.188)	0.577
Communication (district coverage of available telecomm services and networks)	0.139 (-0.205 – 0.483)	0.41
(households affected by Natural Disasters) Drought	-0.377 (-0.833 – 0.08)	0.101
Floods	-0.092 (-0.894 – 0.71)	0.813
Disease Outbreaks	0.038 (-0.808 – 0.884)	0.927
Famine	-0.406 (-0.881 – 0.687)	0.09

* FANCcov is the overall coverage score of FANC programme calculated as the mean of antenatal care coverage, tetanus toxoid coverage, and HIV screening coverage. $^{\beta}$ *FANCscore* is the mean of all the six component scores as applied to FANC programme.

Table 8.2: Univariable analysis for associations between the Overall EmOC Coverage and the EmOC Components & Contextual Variables in 23 districts

Predictor variables	Slope (95% CI)	P-value
Programme implementation strength components	EmOCcov*	
Human resources	0.553 (0.217 – 0.889)	0.003
Service Delivery	0.787 (0.515 – 1.06)	<0.001
Drugs and Supplies	1.014 (0.32 – 1.708)	0.006
Health Financing	0.606 (0.296 – 0.916)	0.001
Health Information System	0.061 (-0.839 – 0.961)	0.889
Leadership and Governance	0.078 (-0.526 – 0.684)	0.79
Overall EmOC Implementation Strength Score Model: EmOCcov = 0.064 + 1.305*EmOCscore [β]	1.305 (0.747 – 1.863)	<0.001
Contextual factors		
Total Fertility Rate	-0.062 (-0.118 – -0.006)	0.032
Infant Mortality Rate	-0.0001 (-0.004 – 0.004)	0.939
Under five Mortality Rate	-0.0001 (-0.002 – 0.003)	0.929
Female literacy levels (adult, primary and secondary education)	0.514 (-2.96 – 3.988)	0.761
Water (hholds using drinking water from improved water sources)	0.763 (0.41 – 1.116)	<0.001
Sanitation (households using improved toilet facilities)	0.759 (0.478 – 1.041)	<0.001
Transport (accessibility to public transport services within 3 kilometres; pass-ability of roads in rainy season)	-0.073 (-0.832 – 0.686)	0.844
Communication (district coverage of available telecom services and networks)	0.189 (-0.079 – 0.458)	0.158
(households affected by Natural Disasters)	-0.209 (-0.591 – 0.173)	0.267
Drought		
Floods	0.305 (-0.328 – 0.939)	0.328
Disease Outbreaks	-0.018 (-0.701 – 0.666)	0.958
Famine	-0.368 (-0.744 – 0.008)	0.055

* EmOCcov is the overall coverage score of EmOC programme calculated as the mean of EmOC availability, institutional delivery rate, and Caesarean section rates.
[β] *EmOCscore* is the mean of all the six component scores as applied to EmOC programme.

8.3.2 Multivariable analysis

Table 8.3 shows the results of multivariable analysis for the overall coverage scores of FANC and EmOC programmes and their respective predictor variables that showed significant association during stage one of univariable analysis. For FANC programme, female literacy was the only predictor variable which was significantly associated with the overall FANC programme coverage (adjusted coefficient=4.798, P=0.011) whereas for EmOC programme, two predictor variables were significantly associated with the overall EmOC programme coverage. These were: EmOC service delivery (adjusted coefficient=0.47, P=0.025) and sanitation (households using improved toilet facilities; adjusted coefficient=0.388, P=0.038).

Table 8.3: Multivariable analysis between overall programme coverage scores and associated predictor variables

Predictor variables	Overall FANC Coverage (FANCcov)		Overall EmOC Coverage (EmOCcov)	
	Unadjusted	Adjusted (95% CI)	Unadjusted	Adjusted (95% CI)
Human resources	0.614	0.036 (-0.544, 0.616)	0.553	0.026 (-0.281, 0.332)
FANC Service	0.83	0.259 (-0.499, 1.017	-	-
EmOC Service	-	-	0.787	0.47 (0.07, 0.87)[β]
EmOC Drugs & Supplies	-	-	1.014	0.037 (-0.712, 0.787)
Health Financing	0.484	0.207 (-0.326, 0.739)	0.606	0.133 (-0.256, 0.523)
Female Literacy	5.501	4.798 (1.275, 8.322)[γ]	-	-
Total Fertility Rate	-	-	-0.062	0.023 (-0.027, 0.074)
Water	-	-	0.763	0.13 (-0.273, 0.534)
Sanitation	0.551	0.262 (-0.244, 0.217)	0.759	0.388 (0.024, 0.752)[ζ]
Famine	-	-	0.368	-0.007 (-0.276, 0.263)

[γ] P=0.011; [β] P=0.025; [ζ] P=0.038

8.3.3 Dose-response analysis (composite scores)

Figure 8.1 shows the dose-response results from the regression line:

FANCcov = 0.358 + 1.007*FANCscore

that involved the overall FANC programme coverage (FANCcov as the response variable) and the overall average score of all the implementation strength components for FANC (FANCscore as the predictor variable). Likewise, Figure 8.2 shows the dose-response results for the regression line:

EmOCcov = 0.0644 + 1.305*EmOCscore

involving the overall EmOC programme coverage (EmOCcov as the response variable) and the overall average score of all the implementation strength components for EmOC (EmOCscore as the predictor variable).

Figure 8.1 shows that, if the overall FANC implementation strength would increase by 1%, the overall FANC programme coverage would equally increase by 1% whereas each 1% increase in the overall EmOC implementation strength would result in about 1.3% increase in the overall EmOC programme coverage (Figures 8.2). The two figures also show that the R-squared value for EmOC programme (=53.0%) was about twice as large compared to that of FANC programme (=25.8%) indicating that the fitted univariable regression model for EmOC programme explained much more of the variation in the overall EmOC programme coverage than the variation explained by the fitted model on FANC programme coverage.

8.4 Discussion

This chapter's main objective was to explore presence and the nature of associations between the coverage and the implementation of FANC and EmOC programmes. Also, in the chapter, a further illustrative dose-response analysis was conducted using univariable and multivariable regressions for FANC and EmOC programme coverages with both the implementation strength components and the contextual factors as explanatory variables. In univariable regressions, the study found that

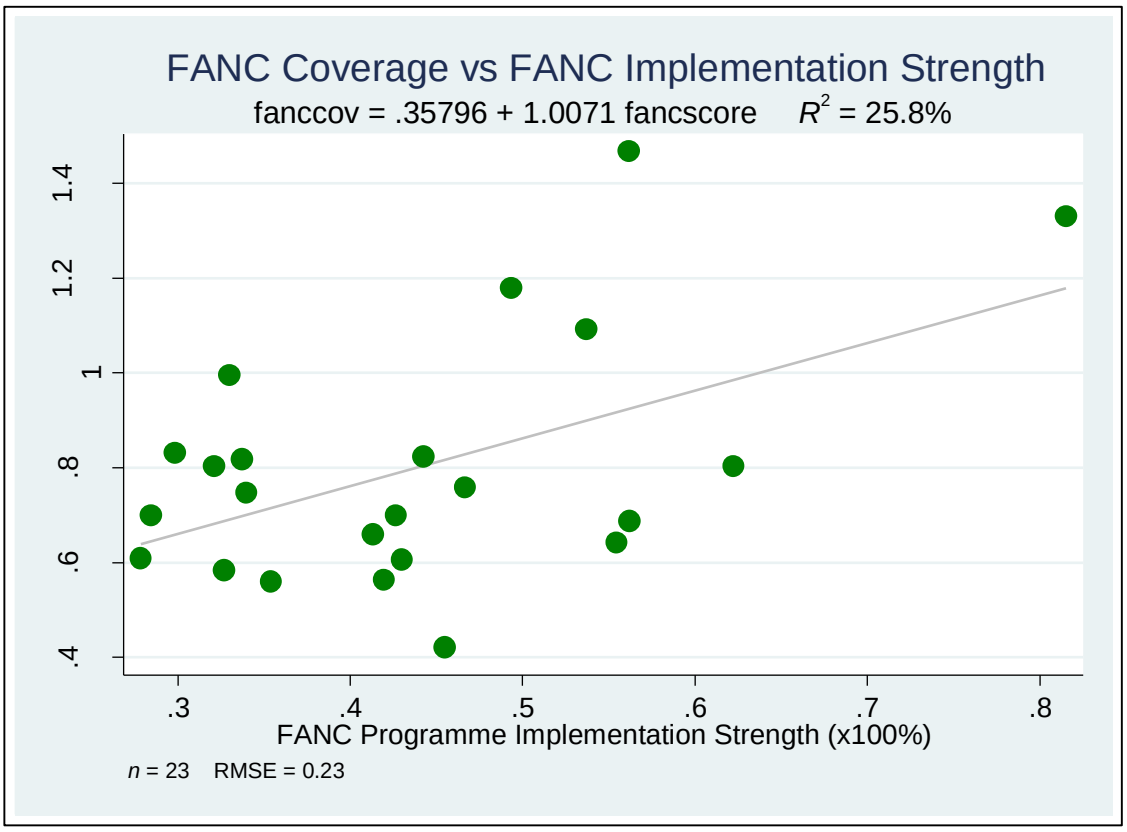


Figure 8.1: Dose-response between FANC coverage and FANC strength
Note: Four data points exceeded 100% FANC coverage and were those for Ilala (147%), Arusha Urban (133%), Songea Urban (118%), and Temeke (109%).[m]

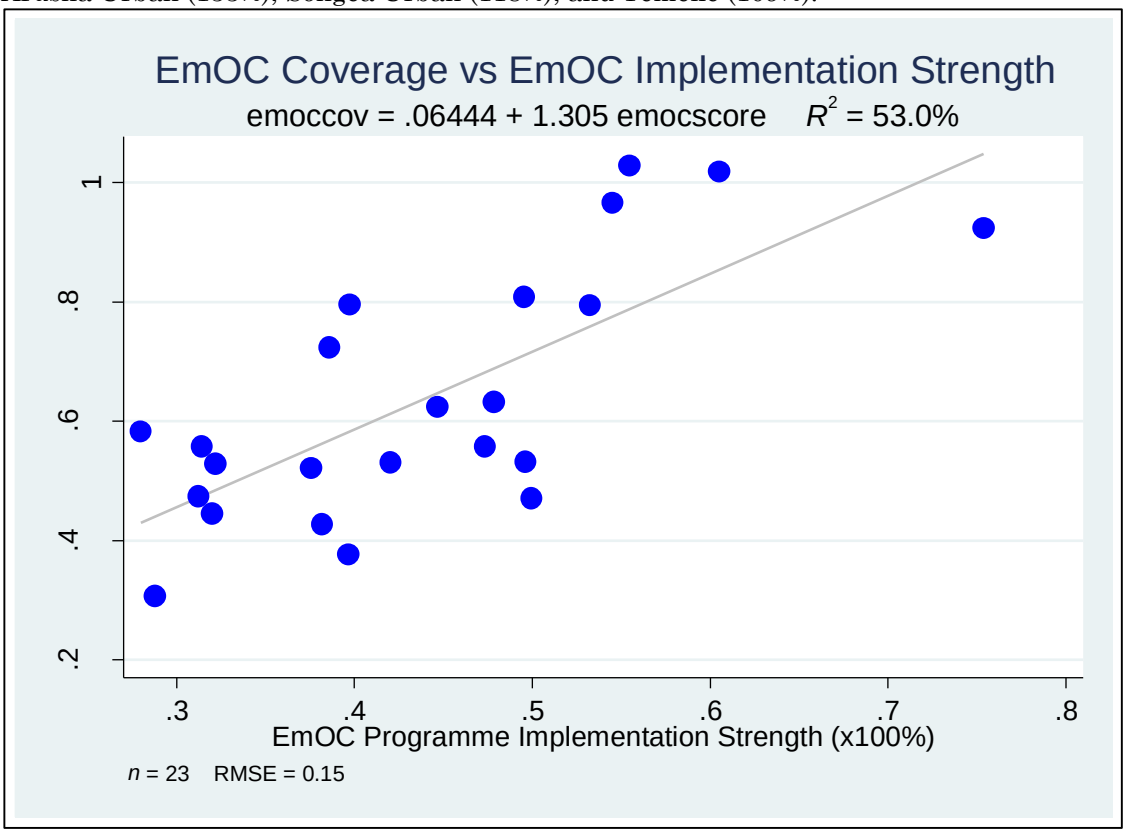


Figure 8.2: Dose-response between EmOC coverage and EmOC strength
Note: Songea Urban and Iringa Urban had coverage of 103% and 102% respectively.[m]

[m] Reasons for coverage exceeding 100% are given on footnote 'j' and in Sections 6.4 and 10.7.

four of the six programme components were linearly associated with the coverage of FANC and EmOC programmes. FANC programme coverage was significantly associated with the human resources, FANC service delivery, and health financing components whereas EmOC programme coverage was significantly associated with the human resources, EmOC service delivery, EmOC drugs and supplies, and health financing components.

Two dose-response models were used to explain the linear relationships between the overall indices for FANC and EmOC programme coverage with their respective overall indices for the implementation strength. The model for the EmOC programme coverage expained twice as much of the variation as the model for the FANC programme coverage – indicating that the EmOC programme coverage was more closely related to the overall EmOC programme implementation strength index than the FANC programme coverage was to its overall FANC programme implementation strength. The illustrative multivariable analysis found that, EmOC service delivery was the only programme component that was significantly associated with EmOC programme coverage after controlling for the four contextual factors associated with EmOC – one of which, was also significantly associatied with the overall EmOC programme coverage (sanitation, adjusted coefficient = 0.388, P=0.038).

The dose-response design is best suited to a situation where there is before-and-after coverage data as well as before-and-after contextual data. In such a situation, it is therefore possible to use multiple regression analysis to explore the relationship between programme implementation strength and programme coverage change, while adjusting for confounders. Given that the study's data on programme coverage and contextual factors were both limited to a single observation time point, results on multiple regression analysis presented here should only serve as a demonstration of the statistical capability for the proposed approach. To exploit such statistical analyses, it is therefore recommendable that future studies should collect data for at least three years for which likely changes can be detected in both programme coverage and contextual factors indicators.

There are limited number of large-scale maternal health programmes that have evaluated programme impact in quantitative measures, let alone having applied dose-response analyis in estimating the impact of programme activities. One rare

example of such an evaluation is the Maternal and Neonatal Programme Effort Index (MNPI) study.[107] The MNPI study developed an index for measuring programme efforts of national-level maternal health programmes, which was related to the approach for this thesis in that they both used maternal health experts. While the MNPI study used the experts to collect its data from the 49 countries, the approach developed in this thesis only used the experts for developing the weighting scales. The MNPI study did not explore the dose-response relationship between programm efforts and programme results, although the average scores across the variables it measured were reported to have risen by 10% in the three years of the study.[107]

Recently, a maternal health programme in Ethiopia looked at the effect of using health extension workers to improve maternal and newborn health care practices and used dose-response analyis to determine the magnitude of the effect.[247] The study found that the response variable of programme intensity had increased by 2.4 times during the two years of the programme in respective communities. Several other studies have used dose-response analyses in assessing the effect of specific interventions on specific maternal health related outcomes.[248] [249] [250]

8.3 Limitations

Data used in these dose-response analyses were only from a single point in time, that is, they were cross-sectional data collected between 2010 and 2011. A single cross-sectional study cannot establish strong associations for which changes over time in both outcomes and contextual factors would be evident, and the impact of programme implementation is likely to be cumulative in nature.[251] Not all indicators of the implementation strength, programme results, and contextual factors had data for both years 2010 and 2011. Furthermore, coverages of both FANC and EmOC programmes included only three indicators for each programme: antenatal care coverage, coverage among pregnant women against tetanus, and on HIV screening – for FANC programme; and availability of EmOC services, institutional deliveries, and Caesarean section rates – for EmOC programme. It is certainly possible that results would have shown a different pattern if each programme had included more indicators in the analysis.

Results presented here were therefore meant to illustrate the approach in evaluating maternal health programmes in low- and middle-income countries as well as demonstrate the analytical process and the likely patterns of associations between FANC and EmOC programme results and programme components and contextual factors. These dose-response results should thus not be used to show implications for making policy. Regarding the sample size in dose-response analyses, designers of the national-evaluation platform recommended that the number of districts be considered depending on "how much the implementation score varies across the different districts, how much variability there is in baseline coverage, and how strongly implementation affects coverage".[66]

8.4 Conclusion

The results of univariable dose-response analyses indicate that FANC programme coverage was significantly associated with human resources, FANC service delivery, and health financing components of programme implementation strength, whereas EmOC programme component was significantly associated with human resources, EmOC service delivery, EmOC drugs and supplies, and health financing components of programme implementation strength. Both health information system and leadership and governance were not associated with the overall coverage of either programme.

In order to effectively benefit from dose-response analyses and generate more and concrete evidence for policy formulation, more data are needed over multiple periods (of at least 3 years) including the minimum essential set of indicators for both FANC and EmOC programmes and all possible contextual factors related to programmes delivery. Resources permitting, the date should include additional indicators on women's personal attributes such as knowledge, attitude and behaviours in relation to FANC and EmOC care services for better understanding of the underlying factors other than programme-related contextual factors.

Part IV: Validity and General Discussion

Chapter 9: Content and face validity of the approach

9.1 Introduction

A few studies have measured the strength of maternal and neonatal health programmes implemented at large-scale using different approaches. For example, in 2005, Ross and Begala[189] reported findings of the Maternal and Neonatal Program Effort Index that was implemented in 1999 and in 2002 in different countries and averaged scores obtained from expert reviewers using a tool with several components. Similar to the maternal and neonatal program effort index study, two other national-level programmes on family planning (1999, 2004, and 2009)[252] and HIV/AIDS (2003)[253] also used average scores, obtained from expert reviewers using tools with several components. Likewise, in 1993, Gold et al reported their findings using scores generated from specific measures combined in four different aspects and with data from different sources.[254] Some of these tools have been used in numerous low- and middle-income countries to report on composite scores that reflected levels of success in implementing maternal and family health programmes.[189] However, evidence from the literature on assessing content or face validity of the instruments used to generate programme data from such large-scale maternal health programmes, is still lacking. Without knowing the validity of instruments used, the overall suitability and acceptability of such tools among users can be hard to establish.

Gaber and Gaber[255] defined content validity as "the fit of the defined content of variables to what is known about the research topic" and face validity as the "common sense appearance of validity in the research findings". Content validity is mainly concerned with the adequacy of incorporating the contents covered by the research tools. Face validity, on the other hand, seeks to answer the question on whether a research makes sense "on the face of things".[256 257] While the two validity measures have been widely reported in studies on newly developed or revised research instruments, face validity has been criticised for having "little or nothing to do with scientific validity evidence"[258] and that it contrasts sharply from scientific

validity evidence.[259 260] Proponents of face validity argue that the central idea behind assessing face validity is "its big picture assessment" and they agree that face validity should not be the only measure of assessment but be "preceded by tests of internal and external validity".[255 261 262]

In recognition of the complex nature and the challenges of implementing maternal health programmes, it is therefore important to have suitable tools for evaluating active programme components responsible for delivering desired impact. Such tools need to be both reliable and valid for measuring the implementation strength of programmes but also acceptable to national and district authorities involved in planning and executing the maternal health programmes. The main objective of the study presented in this chapter was to evaluate the face and content validity of the questionnaire (the Maternal Health Experts tool), which was used to estimate the implementation strength of FANC and EmOC programmes in Tanzania. This chapter assessed the suitability of the questionnaire among maternal health experts, both at national, district and health facility levels.

9.2 Methods

9.2.1 Recruitment of participants

The sampling frame of study participants was all 210 maternal health experts who rated components of FANC and EmOC programmes with regard to their contributing strength to the implementation of maternal health programmes in Tanzania. The 210-participant sample comprised 27 national-level, 28 district-level, and 155 health facility-level participants (as explained in Chapter 3, Sections 3.2 and 3.3).

A convenience sample was selected from the sampling frame. Inclusion criteria included all 27 national-level participants; all district-level participants from six (of the 23) districts; and all staff in charge of health facilities from the six selected districts. The six districts were selected based on the 'activeness' of the district coordinators who would subsequently follow-up to collect questionnaires from district-level and health facility-level participants. All participants had already given their consent to participate in the survey and their signed consent forms had been collected from the survey conducted in May 2012.

All selected participants received three forms: **1: "Maternal Health Experts Tool"** (Appendix 1)**.** This was the main tool that participants had used in May 2012 for rating the strength of FANC/EmOC programme components. The tool was specifically sent back to participants as a reminder of the six components including content and tool layout. The six programme components included: Human resources; Drugs and Supplies; Service Delivery; Health Financing; Health Information System and Leadership/Governance. **2: a 1-page feedback sheet** with a table and graphs showing a summary of the results from the May 2012 survey that participants had participated in; and **3: the survey instrument for assessing content and face validity of the strength tool** (Appendix 4)**.**

For national-level participants, the author sent an email asking each of them to provide their assessment of the tool. National-level participants responded to questions through a website. Telephone conversations and reminder emails were sent for follow-up and to provide clarifications. For district- and health facility-level participants, an email was sent to the six district coordinators with all three forms attached. District coordinators printed each form and distributed them to district-level participants and health facility-level participants. Participants self-administered their assessment of the strength tool on the paper version of the survey instrument. District coordinators collected all paper versions and entered the data through the link that pointed them to the website with the survey assessment instrument. All the data were recorded in a standardized data extraction tool (Survey Monkey) and transferred to Stata v13 for analysis. Table 9.1 shows a summary of participants and the number expected to complete the assessment.

9.2.2 Content and face validity testing

Content validity was assessed through expert review and face validity through review by potential users of the programme strength tool. It has been argued elsewhere that inclusion of experts in the review of contents adds accuracy to the exercise.[263] The study used a reversed 7-point Likert rating scale[264] with 1 = "strongly disagree" indicating lack of agreement, 4 "neutral – neither agree nor disagree" and 7 = "strongly agree" indicating excellent agreement. The 7-point Likert rating scale was considered a fair representation of the variability of respondents' ratings. It has been reported elsewhere that 'more response alternatives will be more reliable than those with fewer' and that '7-point response scales seem preferable to short ones'.[265]

Table 9.1: Number expected and level of participants for the assessment survey

No	Level	Participants	Number expected
1	National	National-level Maternal Health Programme Directors, Managers and Coordinators	10
2	District	The District Medical Officer, the District Health Secretary, the District Pharmacist, the District Reproductive and Child Health Coordinator, the District Nursing Officer	18 (about 3 participants from each of the 6 districts)
3	Health Facility	Staff in-charge of the following health facilities: Uyui District: 1. Ikongoro Dispensary; 2. Isikizya Dispensary; 3. Kanyenye Dispensary; 4. Magiri Dispensary; 5. Mayombo Dispensary; 6. Upuge/Ipuge Health Centre Iringa Urban District: 1. Igumbilo Dispensary; 2. IMECC Hospital; 3. Ngome Health Centre; 4. Polisi Dispensary; 5. Regional Hospital; 6. Sabasaba Dispensary Sumbawanga Rural District: 1. Matai Health Centre; 2. Mtowisa Health Centre Muleba District: 1. Ilemera Dispensary; 2. Kagondo Hospital; 3. Kaigara Health Centre; 4. Kimeya Health Centre; 5. Rubya Hospital Musoma Rural District: 1. Bethsaida Health Centre; 2. Butiama Health Centre; 3. Iringo Dispensary; 4. Kamnyonge Dispensary; 5. Police Dispensary Kilosa District: 1. Kilosa District Hospital; 2. Ilonga Dispensary; 3. Kimamba Health Centre; 4. Ulaya Health Centre; 5. Msowero Dispensary; 6. Mvomero Dispensary; 7. Mvumi Dispensary; 8. Rudewa Dispensary	32 (1 participant from each facility)
Total expected			60

Participants were asked to rate the programme strength tool based on five content validity parameters: 1) overall algorithm 2) tool thoroughness 3) clarity of instructions 4) tool length and 5) tool scoring system (Table 9.2 shows content and face validity parameters measured). Regarding the tool's overall algorithm, participants were asked to rate whether the algorithm was easy to understand and follow. Regarding thoroughness, they were asked whether the programme components comprised the right elements/activities whose perceived effectiveness contributed to the strength of the component (an opportunity was provided for participants to write down any element/activity that they thought was not included under a particular component). On clarity, participants were asked whether the instructions provided were clear and understandable. On the tool's length, they were asked whether the length was reasonable. Moreover, on the scoring system, participants were asked whether the scoring system and interpretation of the scores was clear.

Parameters for testing the programme strength tool's face validity were included in form of two questions: the first question asked participants to rate how suitable the tool was for use in assessing the implementation strength of maternal health programmes, and the second question asked participants of their willingness to use the tool if they had the option to use it themselves. The last section of the tool provided participants with an opportunity to comment with open-ended suggestions that would help improve the tool. For each element asked on content or face validity, an additional response "Don't Know" was provided for participants who were uncertain of their position/response. Basic demographics of participants were also collected, on their: job titles, years of work experience, types of place of work (public, non-governmental or private), education level, age group and gender.

Table 9.2: Content and Face validity parameters assessed

Parameter assessed	7= Strongly agree	6=Moderately agree	5=Slightly agree	4=Neither agree nor disagree	3=Slightly disagree	2=Moderately disagree	1=Strongly disagree	Don't Know/ Unsure
Overall Tool Algorithm [Content Validity]								
The overall algorithm of the tool is easy to understand and follow								
FANC measurement [Content Validity]								
Each component has just about the right elements								
The tool's algorithm for FANC programme is easy to understand and follow								
EmOC measurement [Content Validity]								
Each component has just about the right elements								
The tool's algorithm for EmOC programme is easy to understand and follow								
Overall tool assessment [Content Validity]								
The instructions are easy to follow (understandable)								
The length of the tool is reasonable								
The scoring system and interpretation of the scores is clear								
Overall tool assessment [Face Validity]								
The tool is suitable for use in assessing strength of implementation								
I would use the tool myself								

9.2.3 Analytical approach

Stata version 13 was used to summarize demographic characteristics of participants (gender, administrative level, education level, age groups, years of work experience, and employer types) in percentages. For each item rated by participants, the average score (content or face validity index (VI_{item})) was calculated in Microsoft Excel 365 (Microsoft Corporation, 2013) and rescaled to 0% - 100%. The item's minimum, maximum, and standard deviation values were also calculated and presented with the averages. Validity item scores were calculated according to the formula used by Polit et al as follows: [266]

$$VI_{Item} = \frac{number\ of\ raters\ scoring\ an\ item\ with\ 1, 2, 3, 4, 5, 6\ or\ 7}{Total\ number\ of\ experts}\ \text{X}\ 100\%$$

As responses were ordered (1 – 7), weighted Kappa (κ)[267] was used to assess the agreement among the raters in their three groups of national-level, district-level, and facility-level. Because there were more than two raters (n=42), 10 subjects, and more than two ratings (7 Likert points), the following Stata formula was used, also illustrated in the Stata help file as Example 9:[268]

```
kap rater1-ratern
```

where rater1-ratern indicated a range of raters from rater number one to rater number n, with all raters arranged in columns and subjects as variables (in rows). The Stata command was run three separate times, first time for content validity subjects, second time for face validity subjects, and third time for all subjects. Content validity subjects appear in Table 9.2 and include all but the last two subjects – which were for face validity. Interpretation of the κ scores was based on Landis and Koch criteria for the strength of agreement: ≤0=poor, 0.01–0.20=slight, 0.21–0.40=fair, 0.41–0.60=moderate, 0.61–0.80=substantial, and 0.81–1=almost perfect.[269]

9.3 Results

Table 9.3 shows the demographic characteristics of study participants. Sixty participants were asked to participate, but 42 (70%) completed the survey. 27 (64%) of the participants were female. Seven participants were national-level, 15 district-level and 20 health facility-level.

Table 9.3: Demographic characteristics of participants (n=42)

Personal characteristic	Number (%) of participants
Gender	
Female	27 (64%)
Male	15 (36%)
Level (health system)	
National	7 (16.7%)
District	15 (35.7%)
Health facility	20 (47.6%)
Education level	
Certificate	7 (16.7%)
Diploma	15 (35.7%)
Advanced Dip/Degree	12 (28.6%)
Master or above	8 (19%)
Years of work experience	
<5 years	10 (23.8%)
5 – 15 years	12 (28.6%)
16 – 25 years	12 (28.6%)
26 or more years	8 (19%)
Age group	
<25 years	0 (0%)
25 – 40 years	20 (48%)
41 – 55 yeas	19 (45%)
56 years or above	3 (7%)
Place of work	
Public	28 (67%)
NGO	12 (29%)
Private	1 (2%)
Other	1 (2%)

Most participants had diploma or advanced or degree-level education (27, 64%) with others having a college certificate (17%) and a master's qualification or above (19%). Less than one-quarter of the participants had less than five years of work experience. Other participants had between 5 and 15 years (29%), between 16 and 25 years (29%) and 8 of them (19%) had 26 or more years of experience. Nearly all (93%) participants were aged between 25 years and 55 years. Two-thirds of participants were government employees while less than a third (29%) worked with non-governmental organizations.

Graphs appearing in Figures 9.1 and 9.2 show the percentage scored by each of the 7 Likert scales on the five constructs of content validity and two constructs of face validity. In addition, Table 9.4 summarises the means and standard deviations for each of the 7 Likert scale categories including the minimum and maximum scale categories scored by participants on each of the content and face validity constructs. All of the five constructs of content validity (overall algorithm of the tool, the tool's thoroughness, clarity of instructions provided in the tool, the tool's length, and the tool's scoring system) scored above 88%, with the assumption that raters' agreement was based on the three Likert scale categories of 'slightly agree', 'moderately agree', and 'strongly agree' categories.

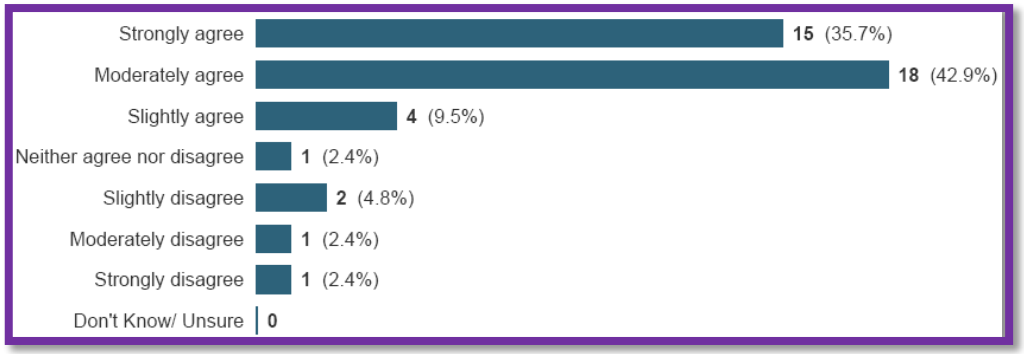
Overall algorithm of the tool received the highest score (92.9%) while FANC and EmOC algorithm section lay outs received the minimum scores each with 88.1%. while raters scored 92.5% and 90.5% on overall contents of the FANC and EmOC programmes respectively, they scored 90.5%, 92.8%, and 90.4% on the tool's clarity of instructions, the tool's length, and the scoring system and interpretation of the scores respectively. Face validity questions on how suitable the tool was for use in assessing the implementation strength of maternal health programmes and on whether raters would be willing to use the tool themselves if they had the option to use it scored 90.5% and 83.4% respectively.

Table 9.4 on the other hand shows that all 10 validity constructs were highly scored with an average of 5.9 "moderate agreement" ranging from an average score of 5.6 (for FANC programme component having just about the right elements/activities) and 6.1 (instructions being easy to follow). The table also shows that, two raters indicated they were unsure or did not know what to rate on whether each of the FANC programme components had the right elements/activities and one rater was unsure or did not know what to rate on the overall algorithm of the tool was easy to understand and follow.

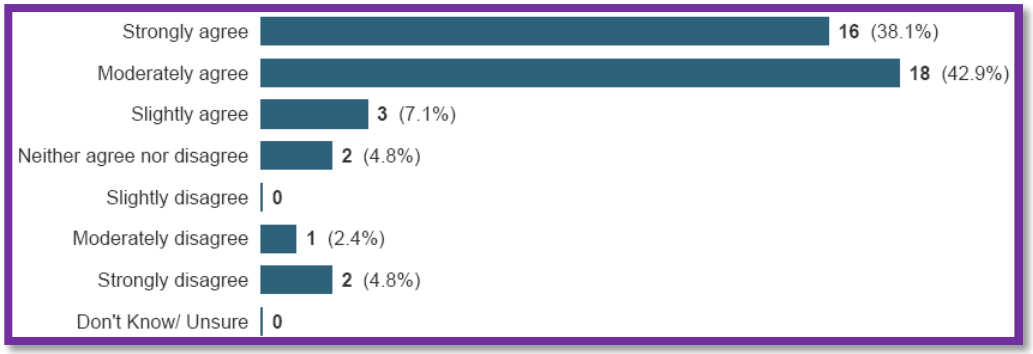
Figure 9.1: Graphs showing 7 Likert scale percentage scores of content validity items (numbers 1 through 4) on algorithm, content and instructions

1. The tool's algorithm programme was easy to understand and follow

FANC

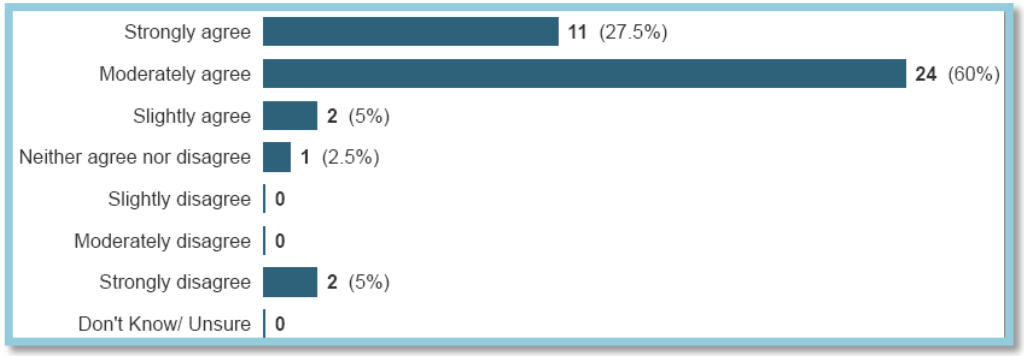


EmOC

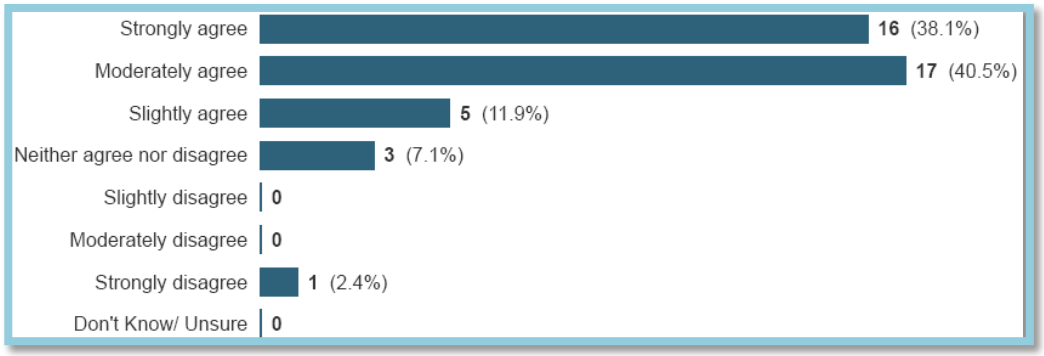


2. Each programme component had just about the right elements/activities

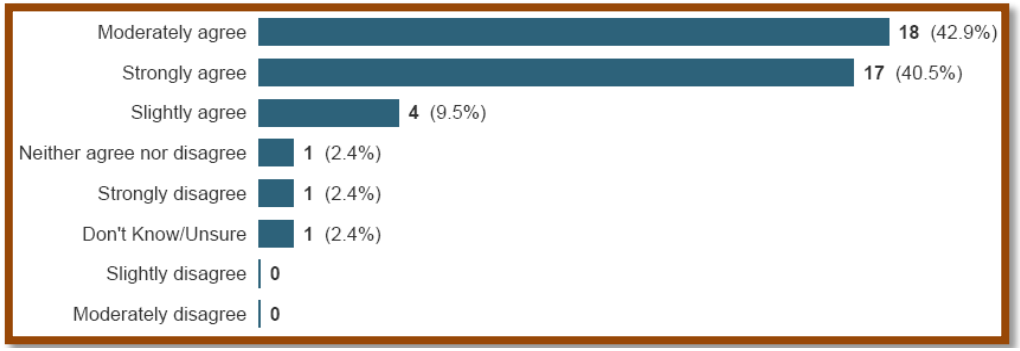
FANC



EmOC



3. The overall algorithm of the tool was easy to understand and follow



4. The instructions were easy to follow

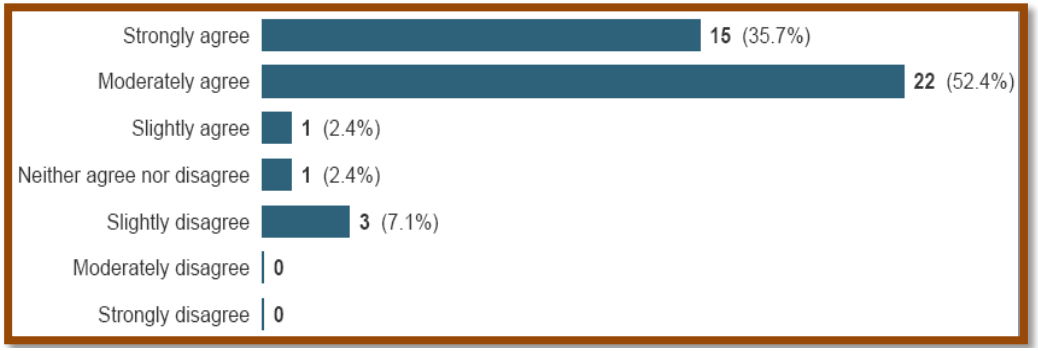


Figure 9.2: Graphs showing the 7 Likert scale percentage scores of content validity items (number 5 and 6) and face validity items (number 7 and 8)

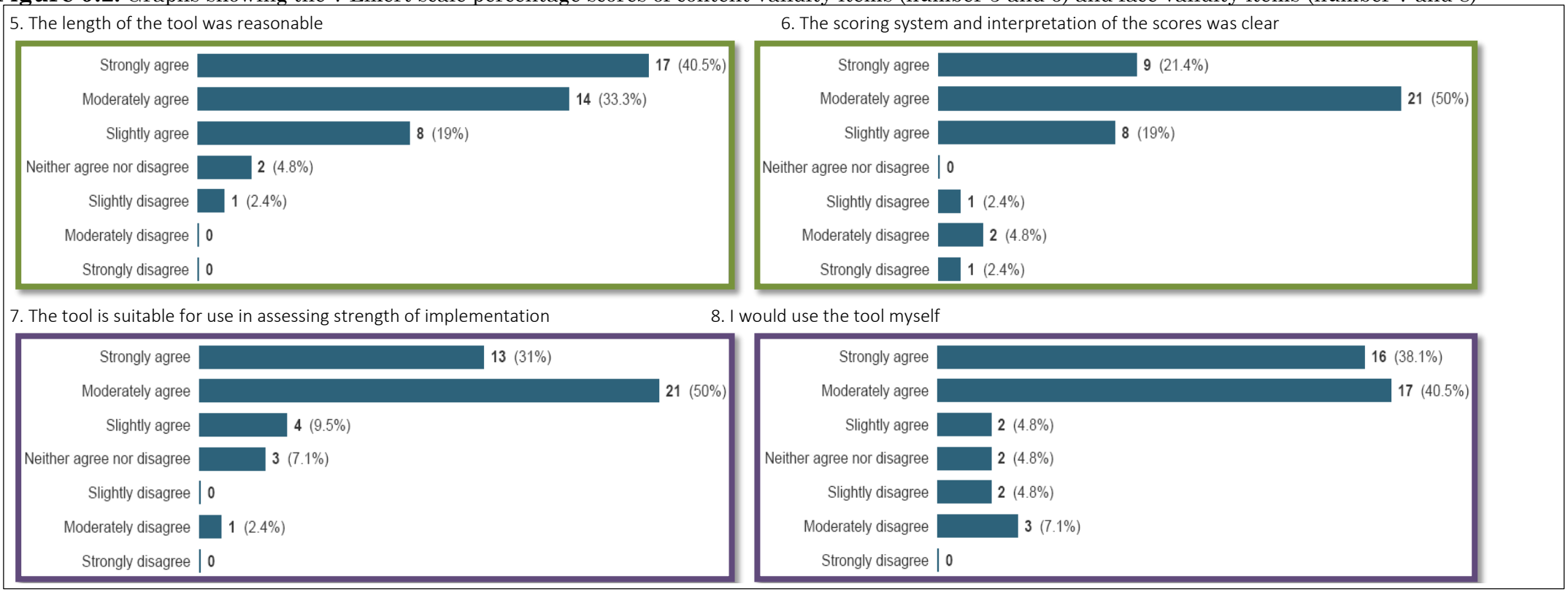


Table 9.4: Results of content and face validity of the implementation strength tool (number of times each Likert item was scored, mean, std dev, min and max)

No	Validity Item	Likert scale (number of times scored)								Mean	Std Dev	Min	Max
		0	1	2	3	4	5	6	7				
1	The tool's algorithm for FANC programme is easy to understand and follow	-	1	1	2	1	4	18	15	5.9	1.4	1	7
2	Each FANC programme component has just about the right elements/activities	2	2	-	-	1	2	24	11	5.6	1.8	0	7
3	The tool's algorithm for EmOC programme is easy to understand and follow	-	2	1	-	2	3	18	16	5.9	1.5	1	7
4	Each EmOC programme component has just about the right elements/activities	-	1	-	-	3	5	17	16	6.0	1.2	1	7
5	The overall algorithm of the tool was easy to understand and follow	1	1	-	-	1	3	18	18	6.0	1.4	0	7
6	The instructions are easy to follow (understandable)	-	-	-	3	1	1	22	15	6.1	1.1	3	7
7	The length of the tool is reasonable	-	-	-	1	2	8	14	17	6.0	1.0	3	7
8	The scoring system and interpretation of the scores is clear	-	1	2	1	-	8	21	9	5.6	1.4	1	7
9	The tool is suitable for use in assessing strength of implementation	-	-	1	-	3	4	21	13	6.0	1.0	2	7
10	I would use the tool myself	-	-	3	2	2	2	17	16	5.8	1.5	2	7

Note: Likert scale categories: 0= Don't Know/ Unsure; 1=Strongly disagree; 2=Moderately disagree; 3=Slightly disagree; 4=Neither agree nor disagree; 5=Slightly agree; 6=Moderately agree; 7= Strongly agree

Results for weighted kappa agreement of the raters within their three system level-groups are as shown in Tables 9.5, Table 9.6, and Table 9.7 for content validity, face validity, and both content and face validity subjects respectively. All three tables, Tables 9.5, 9.6, and 9.7 show two findings: first, with the exception of national-level raters based on both content and face validity (Table 9.7), all the weighted Kappa scores ranged from between less than 0 and 0.2 under Landis and Koch criteria of poor and slight agreement among raters. The weighted Kappa score for the national-level raters on content and face validity was about 0.022 meaning the strength of agreement among them was fair under Landis and Koch criteria. The second finding on the tables is that none of the weighted Kappa scores were significant, with P-values ranging from about 0.3 to 0.88.

Table 9.5: Weighted Kappa scores for content validity

Level	Kappa	Z	Prob>Z
All	-0.0013	-0.17	0.5690
District	-0.0098	-0.44	0.6717
Facility	-0.0178	-1.05	0.8524
National	-0.0305	-0.63	0.7354

Table 9.6: Weighted Kappa scores for face validity

Level	Kappa	Z	Prob>Z
All	-0.0161	-1.05	0.8530
District	-0.0533	-1.18	0.8815
Facility	-0.0122	-0.34	0.6325
National	-0.0868	-1.02	0.8451

Table 9.7: Weighted Kappa scores for both content and face validity

Level	Kappa	Z	Prob>Z
All	-0.0034	-0.50	0.6915
District	-0.0121	-0.62	0.7321
Facility	-0.0146	-0.95	0.8293
National	0.0218	0.53	0.2996

9.4 Discussion

This study has explored the content and face validity of a new tool proposed for helping in planning and evaluating FANC and EmOC programmes in Tanzania among those responsible for national coordination and district planning and execution of programmes. It has given information on five constructs (subjects) for evaluating the content of the tool and two constructs for assessing the tool's acceptance and use (face validity). Overall, the raters 'moderately agreed' (at an average of 5.9 on the 7-point scale) on the overall content and face validity of the tool. According to the weighted Kappa statistics, the strength of agreement among the raters was between poor and fair on all subjects of content and face validity.

On content validity:

Using the validity assessment tool (Table 9.2), the raters were required to evaluate the content validity of the implementation strength instrument (Appendix 1: Maternal Health Experts Tool). In evaluating the content validity, the study used four commonly used constructs of: algorithm, thoroughness (completeness), clarity of instructions, and tool's length.[270] [271] [272] The study also assessed the tool's scoring system – as an additional content validity construct. As is usually required, content validity is supposed to be reviewed by experts in the field for which the tool will be used.[273] As was the case, the study involved more than 40 participants from those involved in coordinating and planning FANC and EmOC activities to those who implement them. The raters also had varying backgrounds in education levels with various number of years of work experience, age, and gender and included people from the public and non-governmental sectors. With an overall score of 88%, raters indicated that the designed tool had high content validity.

In rating the five constructs on content validity, raters were provided with space within the rating tool to comment on completeness of the implementation strength instrument. Additional comments provided by participants were aimed at improving the instrument and were centred on adding or removing an element within a specific programme component (Table 9.8). One notable comment from a participant pointed out that the tool "…needs to be applied/used by a team, rather than a single manager", which was in fact the primary aim of the tool. Future use of the tool would therefore benefit from incorporating the comments given by the raters.

Table 9.8: Selected Comments from Participants on each component of the assessment tool

Relevant component	Participants comments
Human resources	"ability to forecast on human resources" should be added as element/activity under human resources for both FANC and EmOC programmes "The quality of pre service training for medical doctors is not on the list, use of standard work and checklist, Quality improvement initiative. There is no standard in clinical officer school, curricula are outdated, no checks on whether competency effected, lack of post training supervision, mentoring/ coaching. Rotating with specialists will not solve this." "some of the sub elements were vague and diluted the bundle weight; Advocacy/policies to be separate from management" "We should not involve Traditional Birth Attendants in giving them skills to perform deliveries since we will increase home delivery but we should use them to escort pregnant mothers to deliver at health facilities" "refresher training courses on EmOC and FANC" should be added as an activity
Service delivery	"Space for ANC consultation" should be added as an element regarding infrastructure within service delivery component
Drugs and supplies	"The ability to forecast Drugs and Supplies" should be added
Health financing	Three participants said that they "did not quite understand the health financing category"

Health info system	None
Leadership/governance	None
Overall Comments	"The questionnaire did not show clear the score agreement and disagreement… not clear to understand its arrangement" "Some of the sub elements were vague and diluted the bundle weight; Advocacy/policies to be separate from management" "The big challenge with this tool is that, the language of the proposed programme elements are too broad which gives room for different interpretation and in some categories there are more than one element e.g. increasing percentage of government health budget spent on (three element have been put in one rank) it will be difficult to have an objective assessment and it will be difficult to know out of the three element which one needs further attention if the score is low. I propose not to have more than one element for the ranking (i.e. do not put in a, b, c or bullets). Also consider having a glossary / definition / or meaning of terms used. e.g. practical training exposure..., development by decentralization...., etc.(this can mean different things by different people who will do the assessment)" "Overall I think the tool is very good but I'm not sure whether any single person that might use it would have all the information they would need to assess each of the components. So it probably needs to be applied/used by a team, rather than a single manager."

On face validity:

The study found that face validity of the tool was rated high with its two constructs scoring an average of 87%. Face validity is based on the argument whether the proposed tool will be relevant for use.[274] [275] With the raters' score on face validity, there is an indication that the proposed tool will be suitable for future use in assessing the implementation strength of maternal health programmes.

On overall validity:

While the percentages reported were encouragingly high, weighted Kappa scores on both content and face validity showed either poor, slight, or fair – reflecting that the strength of agreement among the raters was on average low. Weighted Kappa scores were based on individual Likert points for each construct. Even though the reported strength of agreement among raters was low, all P-values for the weighted Kappa scores were larger than 0.29 meaning that none of the scores was significant and that the precise extent of agreement among the raters is unknown and could have taken any value ranging between 0 (poor) and 1 (almost perfect). The magnitude of Kappa agreement scores could have been affected my several factors – one of which could be the number of constructs. Gwet argues that "a Kappa value of 0.6 based on 80 subjects conveys a much stronger message about the extent of agreement among raters, than a kappa value of 0.6 based on 10 subjects only.[276] Interpretation of the scores can therefore mean differently depending on the context. The low weighted Kappa scores could have also being affected by the number of Likert points on which raters used in rating the constructs. It is possible that, if raters could rate a construct based on a limited number of points such as on a Yes/No scale, weighted Kappa scores would have showed much higher magnitude of strength in agreement among the raters. A similar finding was reported in an experiment which varied categories (cf Likert scales used in this study) from 2 to 5 and subjects from 2 to 60 and noted that Kappa scores decreased as "the number of subjects or the number of categories increased".[277]

Applicability of the proposed tool:

The goal of this study was to evaluate the instrument's suitability in evaluating FANC and EmOC in Tanzanian districts by assessing whether the instrument was

properly designed (content) and that it will be acceptable among its potential users. While districts in Tanzania regularly collect data (mostly on quarterly basis), evaluation of programmes generating the data need to be closely examined through acceptable tools that have passed the validity test among its users. As revealed in this study, the instrument is more likely to be supported by national coordinators and district planners and programme implementers as the Ministry of Health and Social Welfare endeavours to implement maternal health programmes in districts. While the sample size for this study was not determined using study power calculations, the number of participants included in the study (n=42) was consistent with numerous other studies that assessed content and/or face validity. Few examples of such similar studies include: a study in Kenya used 15 participants to assess the face validity of a tool for measuring pain;[278] an American study had included 38 prospective participants in assessing face, content, and construct validity of a commercial tool on da Vinci Skills Simulator;[279] and, a study on content validity of a tool for measuring diabetes patient-reported outcomes used in-depth interviews with 45 patients and parents/caregivers of patients with Type 1 diabetes;[280]

9.5 Conclusion

The goal of this study was to assess the content and face validity of an instrument (FANC and EmOC programme implementation strength tool) based on perspectives of national-level and district-level maternal health programme coordinators and managers, as well as on the perspectives of staff responsible for day-to-day implementation of maternal health activities in health facilities. The instrument is intended for use by mainly district health managers during preparations of annual comprehensive district health plans. The instrument helps planners in allocating resources based on the proportional contribution of each programme component to the implementation strength of maternal health programmes or interventions within a district. To help improve the instrument and promote its use, a follow up qualitative study involving in-depth interviews with groups of participants is required. This will help incorporate other left-out programmatic approaches, and cultural and professional contexts. More research is needed before this instrument is used with wider populations of potential users. As the study collected ratings from individuals on self-reports, results from focus group and in-depth individual interviews will improve the content of and the face validity of the proposed tool.

Chapter 10: Application of the implementation strength to programme evaluation and maternal health policy

10.1 The approach's validity

10.1.1 The approach and its key results

This thesis has developed and validated a novel approach called "programme implementation strength scoring" designed for use in evaluating large-scale maternal health programmes, especially for low- and middle-income countries. The approach was informed by perspectives and views of Tanzanian policy-makers and implementers. While low- and middle-income countries are scaling up implementation of health programmes with the particular aims of achieving the millennium development goals and other local and global initiatives, rigorous methods are needed to inform implementers and policy makers of the effectiveness of the programme's processes and outcomes. The programme implementation strength approach applied in this investigation (stages of which are shown in Figure 2.2, and presented and discussed in Chapter 2) was guided by a conceptual framework (Figure 1.1). It should also be noted here that, in reference to Victora's national evaluation platform,[66] the proposed approach is designed to use the district as its unit of measurement and analysis.

To recapitulate the implementation strength approach's stages: For the planning stage (the 'P' in the 'PIA' stages refers to planning), the author conducted a review of the literature from which the possible indicators for use in evaluating the implementation strength, programme results, and the contextual factors were identified. This review was followed by the author conducting a survey of a group of maternal health experts in order to obtain their opinions of the proposed tools for use in the evaluation process. The experts' opinions were subsequently used to refine the tool which was followed by using it to collect relevant data from study districts. In addition, the experts' opinion survey informed the weighting scales that were used

to calibrate district programme data. For the information-gathering stage (the 'I' in 'PIA' refers to information gathering), programme monitoring data were retrospectively collected for all proxy indicators of the implementation strength (programme components/the six building blocks), programme results (coverage) as well as those for contextual factors.

For the analysis stage ('A' in 'PIA' refers to analysis), the author analysed all data categories including an illustrative dose-response analysis for the association between programme implementation strength components and programme coverage indicators. With guidance from the OECD handbook (supported by a review of similar guidelines),[100] procedures and analytical methods were followed to ultimately generate the aggregate scores particularly those for the implementation strength. Programme monitoring data for the two tracer programmes and data on contextual factors were obtained from cross-sectional surveys and other data sources. Not all indicators had data for both years (Section 8.1) even though data were collected for the period between January 2010 and December 2011. As a result, this study could not investigate the relationship between implementation strength and changes in coverage over time, which would be needed to generate conclusions regarding programme evaluation.

The key findings obtained from the expert opinions and from data for the implementation strength clearly suggested that all six components were important in contributing to the strength of programmes (Section 5.5). Further, PCA results (in Chapter 5) indicated that the six programme components 'naturally' grouped into two categories: one category formed by the four components of human resources, service delivery, drugs and supplies, and health financing (which had high weightings within the first principal component); and a second category formed by health information systems and governance and leadership (which had high weightings within the second principal component). It was, however, noted that EmOC drugs and supplies also joined the health information system and leadership and governance components with high weightings within the second principal component (Section 5.4.1.2).

Even though all six components contributed to the strength of FANC and EmOC programmes, some of them contributed more to the overall programme strength than others. The human resources component carried more weight than other programme

components (Figures 3.3, 3.7, 5.9 and 5.10). Of the three elements of human resources (availability, competence, and productiveness), experts gave more priority to, availability (referring to having adequate staff: numbers of doctors, nurses and midwives) than competence and productiveness (Tables 3.5 and 3.6). Further reflection on these results appear under Section 10.4 below.

In the literature, programme implementation strength has sometimes been referred to as programme intensity. Programme intensity was used in evaluating maternal health programmes by the Maternal and Neonatal Program effort Index (MNPI) study and in subsequent evaluations conducted in some of the countries where the index was first applied.[97 98 107] The MNPI was itself adapted from earlier versions that assessed the intensity of family planning programmes,[252] and in the last few years, evaluations of HIV/AIDS programmes have used similar approaches.[253] Most recently, implementation strength has been used to assess the intensity of large-scale maternal and child health programmes, such as the studies in Ethiopia by Karim et al[247] and Miller et al[84] and a Malawian study on community case management of childhood illness.[281] Karim et al's study further applied the results of the programme intensity measure to a dose-response analysis.

The implementation strength approach proposed in this thesis adds to previous work in several ways. For example, the MNPI study used expert opinions in rating its 13 programme components but neither collected data on implementation strength, nor did it test the index's validity. In addition, except the Malawian study, none of the other cited studies reported the validity of their approaches or tools used for evaluating their respective programme's implementation strength. The Malawian study later piloted the validity (sensitivity and specificity) of the implementation strength method through cell-phone interviews to health surveillance assistants on key indicators of the programme's implementation strength.[85]

In contrast to the other implementation strength studies, this study developed and tested the approach in evaluating two large-scale maternal health programmes and presented the results of the three stages of the process. Based on content and face validity, the tools used for assessing the implementation strength were rated "very good" overall by the people likely to use the approach in evaluating programmes. Furthermore, looking back at the thesis's first research question: "what are the active components responsible for the implementation of FANC and EmOC

programmes?", by developing the implementation strength approach, the thesis identified and adopted the WHO health system building blocks as integral components well-fitted for implementing large-scale maternal health programmes.

There are few important lessons to draw from the implementation strength approach. The matrix shown in Table 10.1 summarises these by presenting pros and cons in comparison with other evaluation methods. The matrix summarises the theoretical and practical applicability of some common evaluation methods, highlighting key aspects (column 2) and briefly pointing out how each method is applied (columns 3). Column 4 of the matrix shows pros and cons of each method, and was the inspiration for developing the approach for this thesis. The main advantage of the approach is its design flexibility. Depending on the purpose of the planned evaluation, the proposed approach can be moulded to evaluate either or both of the implementation processes or programme outcomes.

While introducing the approach, Section 2.6.1 emphasised that the approach would require continuous monitoring of programme activities and processes, and programme contextual factors as well as programme outcomes to maintain advantages over other evaluation approaches. The section also noted that, in order for the approach to describe the nature of the association between implementation strength and programme results, it was imperative to perform a dose-response analyses with contextual factors as potential confounders. The dose-response analysis was facilitated by the approach's ability to collect indicators on both sides of the equation (indicators for programme coverage on the response side and indicators for the implementation strength and contextual factors on the dose side).

Despite its advantages over other evaluation methods, the implementation strength approach has several limitations. The primary limitation is cost, due to continued data collection over time. Programme monitoring data should be continuously collected over the lifetime of the project, – which can be labour-intensive to the already overworked health workers especially in LMICs. It can be difficult for the health system to find additional resources to facilitate data collection. However, it is recommended that a minimum set of indicators should be selected jointly by interested parties in order to reduce labour and cost. In evaluating health system-related programmes, most of the selected indicators can be embedded within (or selected from) the family of indicators whose data are part of routine data collection

Table 10.1: Key evaluation approaches for large scale public health implementation, with pros and cons

Evaluation Approach	Key Aspect(s)	How the approach works	Pros and/or cons
Programme Implementation Strength	• Programme Implementation Strength tracks both process and outcome indicators through which evaluators can estimate in statistical terms, an association between programme activities and programme results • The approach also tracks contextual factors (facilitators/inhibitors) which can interact between processes and outcomes and can therefore be treated as confounders	• The implementation strength is conducted in three stages: 1) Conducting a review of the literature and expert opinion survey in order to: identify the essential components of a programme, the contextual factors, likely programme results, and their corresponding proxy indicators; prepare the aggregation, weighting, and scoring procedures and the data analysis plan 2) On continuous basis: quantify/value resources and inputs, monitor programme activities, and, document all identified contextual factors in programme areas (usually districts – for large-scale programmes) 3) Analyse all data and, conduct dose-response analyses to explore associations between strength and results	• By continuous monitoring of programme activities (processes), programme contextual factors and outcomes, implementation strength can be used as a process evaluation method as well as for evaluating programme impact • Through dose-response analyses, the approach can precisely describe the nature of the association between implementation strength and programme results with contextual factors as potential confounders • However, its demand for continuous data collection over time, the approach will require a minimum set of indicators to be economical and require less work on behalf of implementers. Selection of the indicators can attract bias and produce biased results
Process Evaluation	• A process evaluation assesses the course of a programme on how programme activities are delivered and the context of a programme • Process evaluation describes programme activities (types of activities, levels of programme participant participation, and the programme quality	• The following activities can be used as a means to conducting a process evaluation: 1) Determine the type of information required and questions to be answered 2) Select appropriate methods for use in getting the information. Methods can include conducting focus groups, structured interviews, surveys, review of records and/or documentation. 3) Share the results with stakeholders (participants, implementers and funders)	• Good for measuring programme fidelity • Contributes to understanding relationships between activities and outcomes • Useful for assessing accountability and programme improvement and replication • Can be implemented even with limited resource depending on method of inquiry • Process evaluation may not be suitable for impact evaluation resulting in quantitative estimates
Theory of Change	• A Theory of Change is a tool that involves evaluators developing a theory of how the intervention will bring about expected changes in the outcomes • Theory of Change can be used to guide process evaluation	• In order to develop a Theory of Change, five steps are required: [282] 1) Identify the long-term goal of the project/programme 2) Conduct "backwards mapping" to identify the preconditions necessary to achieve that goal 3) Identify the interventions that the initiative will perform to create these preconditions 4) Develop indicators for each precondition that will be used to assess the performance of the interventions 5) Write a narrative that can be used to summarize the challenges in the theory	• Theory of Change improves the design and evaluation process of complex and multi-faceted interventions. • It does not involve estimating a quantitative measure of the programme effects • Even though a useful monitoring and evaluation planning tool, Theory of Change cannot be used to estimate the extent (size) of programme outcomes/impact

Randomised Controlled Trials	• RCTs randomise the allocation of participants receiving the intervention and those who do not	• There are three stages through which an RCT is conducted: 1) Setting up the protocol (including defining: primary and secondary outcomes; the study population; inclusion and exclusion criteria; sample size; a statistical plan; design of case record forms; and likely logistical issues) 2) Conducting the study (including monitoring of the study, data collection, database design and data entry) 3) Conducting statistical analysis and results reporting	• RCTs are considered the 'gold' standard for evaluating outcomes and impacts of interventions • Besides having high internal validity, RCTs have very low external validity especially in large-scale programmes. • RCTs are mainly useful for impact evaluation
Difference-in-Difference (also known as a before/after intervention/comparison design)	• Difference-in-difference method tracks both the intervention and control (comparison) group to determine whether the intervention is associated with any improvement over and above that occurring in the comparison group over time. • Difference-in-difference requires data to be measured at two or more different time periods	• To estimate the difference-in-differences between the intervention group and the comparison group, the average change over time in the selected outcome is calculated: • If data were collected in two time periods only, a normal regression analysis is used (ordinary least-squares regression or logistic regression) • If there are more than 2-time periods, a regression with fixed effects for time and group is applied	• The difference-in-difference method can be useful for both impact evaluation and for process evaluation. However, it can be very challenging to obtain data from a comparison group in situations where there is potential for spill-over effect
Realist Evaluation	• A realist evaluation is a theory-driven type of evaluation. Underlying theories of a programme about the way in which context and mechanism affect outcomes are developed, tested and refined • Realist evaluation asks 'what works for whom in what circumstances and in what respects, and how?'	• To perform a realist evaluation:[76] 1) Draw out and formalise programme theories to be tested by searching through documents, programme planners and implementers, previous evaluation studies and other literature. 2) Collect data (both qualitative and quantitative) to cross-examine initial hypotheses (preliminary theories) on both the Contexts, Mechanisms and Outcomes. 3) Conduct a systematic test using the collected datasets by identifying patterns in the outcome's successes and failures within and across interventions while continually generating and testing hypotheses regarding the implementation 4) Assess and interpret the results. Have the theories about how the programme worked, been supported, or refuted by the analyses? 5) Draw out a 'context-mechanism-outcome' (CMO) framework by filling out each of CMO components and observing the interaction between context and the mechanisms that result in outcomes	• Realist evaluation is best for establishing a theory-based relationship between processes and outcomes • While they are not necessarily intended to evaluate impact at all, realist evaluations are considered to be more methodologically robust than process evaluations (by some) because, other than collecting similar data but with the element of hypothesis testing, with realist evaluations, it is possible to have a more robust means of exploring whether the conclusions drawn are correct • Like process evaluations, realist evaluations are time- and resource-consuming in terms of the amount of data required and the amount of time needed for analysis

systems (HMIS). Care is therefore needed in selecting the indicators – a process that can be biased depending on the political, programmatic, and funding influences of those involved in indicator selection. Because the approach can be used as an accountability tool through ranking districts based on implementation strength scores (such as Table 5.2 and Table 5.3) and programme results (such as Table 6.1 and Table 6.2), the approach can also attract some criticism from district health authorities who might find it less favourable for their efforts. The following sections extend the discussion on the approach's theoretical comparison to realist evaluation (Section 10.1.2) and theory of change evaluation (Section 10.1.3), as well as the approach's potential to capture equity dimensions (Section 10.1.4).

10.1.2 A realist evaluation's view of the implementation strength approach

Realist evaluation is another possible approach that could be used in the conceptual framework for the implementation strength method (as mentioned in Chapter 2). The conceptual framework for the approach could have used realist thinking. Realist evaluation requires evaluators to populate the Context-Mechanism-Outcome framework by formulating programme theories to be tested (about why the programme worked, under what circumstances, and so forth), collecting data, conducting a systematic test, and assessing results of the collected data.[76]

While the conceptual framework for this approach followed the Donabedian approach[131] which was an adaptation of the IHP+ framework,[63] the process undertaken fulfilled several aspects of the CMO framework. Even though the conceptual framework for this thesis (Figure 1.1) had no particular theories formulated for testing, by reviewing the literature and involving the maternal health experts at some level, the process partially framed out a theory that issues from each of the six blocks were responsible for generating the anticipated programme results. The realist approach would have attempted to describe the mechanisms that resulted in the outcomes that were seen, which would have required additional qualitative data collection.

The moderating conditions in realist evaluation are the equivalent of this framework's contextual factors (Section 1.7.4 and Chapter 7). The realist evaluation's underlying mechanisms would be equivalent to outlining the processes

(programme activities) for delivering the FANC and EmOC services to target populations. One other diverging factor of the framework from realist evaluation is that the latter would describe its outcomes in view of how initial theories would have been modified through collected data, and resulting in continually adapted theories.[283]

10.1.3 A theory of change's view of the implementation strength approach

The theory-driven approaches to evaluation are based on understanding the fundamental programme's theory on whether and how the programme works.[284] By understanding the 'theory of how and why an initiative works',[285] a theory of change can lead to a hypothesis of the causal linkages to programme outcomes. Implementation of large-scale programmes is a complex process, and, while a large-scale maternal health programme can be financially demanding and human resource-intensive, organising the resources and implementing programme activities may not automatically result in delivering the services to populations.

The theory of change's view of the conceptual framework would thus start with framing out the desired programme results and working backwards to determine the probable conditions to be fulfilled in order to generate the results. By so doing, a hypothesis on the causal pathways can be generated. A final task required in preparing the theory of change is 'an accompanying narrative that describes the pathways and key assumptions'.[286] In the current study, one hypothesis for the theory-of-change could have been stated as 'increased coverage and utilisation of FANC and EmOC services in study districts is associated with increased implementation of the programmes delivering the services'.

Achieving higher utilisation and coverage levels would have embodied several assumptions such as that, under favourable contextual conditions, FANC and EmOC programme activities were implemented as planned and that outputs and outcomes were produced as anticipated. Examples of the assumptions for the programme activities could have included (but are not limited to): health facilities had an adequate workforce who are competent, responsive and productive (motivated); essential drugs and supplies for antenatal care and emergency obstetric care were well stocked throughout the year; funding to districts for executing maternal health

services was sufficiently allocated in a timely fashion; and there was a large number of geographically distributed and easily accessed facilities. Some of the assumptions for programme outputs and outcomes could include, but are not limited to, pregnant and postpartum women being aware of the services provided and having the will to demand and the ability to access the services in facilities.

It can therefore be generalised here that framing the concept of the approach could have alternatively been done using the realist evaluation or the theory of change as two other possible frameworks. However, the focus of this study was on developing a measure of implementation strength rather than applying realist thinking or theory of change perspectives with the goal of a quantitative estimate of effect of the programme, whereas realist thinking might have shed more light on how the programme effects varied by context, and theory of change could have been useful in developing the study's framework but would not necessarily lead to a dose-response analysis. Based on the theoretical and practical evidence in the validity and relevance of the conceptual framework used, it can be concluded that the applied framework served its purpose well in guiding the implementation strength approach in data collection and data analysis.

10.1.4 Capturing equity dimensions in the implementation strength approach

Through continuous monitoring of programmes, the implementation strength approach has the potential to inform service coverage amongst vulnerable populations. Vulnerable populations for the FANC and EmOC programmes are women of lower socioeconomic status and those based in rural settings that are hard to reach.[130 236] Efforts to add an equity perspective to implementation strength analysis could begin by adding one or two equity indicators to be reported at regular intervals, with specific ongoing analysis being disaggregated by geographically-defined structures such as wards, ethnic groups, or relative poverty in order to track service coverage over time. Examples of some of the equity indicators that can be documented from all first antenatal care clinic attendants could include, but are not to be limited to: number with access to improved toilet facilities; number with access to improved water sources; number with secondary education; and self-reported health status. The number of community health workers in geographically-defined structures such as wards (for example per 1000 members of the population) can be

obtained from the district medical officer. By assessing changes in the proposed equity indicators, routine data can help to initiate appropriate responses from district health authorities by targeting wards, those of low socioeconomic status, or ethnic groups having the poorest equity indicators.[287] Responsive actions for addressing equity can include health system strengthening, for example through equitable distribution of health workers, improved logistics for essential drugs and equipment supply, increased supportive supervision of health facilities within affected catchment populations, and so forth. Other non-health system actions could involve improving social services, geographical infrastructure and other contextual factors discussed in this thesis so that resources are leveraged to improve the health of all populations in the district.

On the other hand, it is also possible that continuous monitoring of routine indicators through the implementation strength approach could potentially blind programme implementers with the aggregate coverage statistics and fail to capture equity disparities within populations.[288] Cross sectional household surveys should therefore be part of efforts to capture equity data (such as socioeconomic indicators), which the approach is otherwise unable to capture through routine indicators. Household surveys have an additional benefit of acquiring more comprehensive and focused information compared to routine monitoring indicators. A good example of a survey that generated equity information is a recent study in Ethiopia, which used a simple tool to capture five equity-related variables (roof material, ownership of a table, type of toilet facility, ownership of a radio and educational attainment) to identify a pattern of EmOC service utilisation in women.[289] Through the demographic and health survey, the study interviewed women being discharged from a hospital's maternity ward and found that women who used the EmOC services were urban residents with lower parity. Another study in 35 Countdown-to-2015 countries used demographic and health surveys or multiple indicator cluster surveys and found that the overall coverage of maternal and child health programmes "were associated with improved equity".[290]

10.2 Potential applicability

The proposed implementation strength approach comes at a time when the Tanzanian Ministry of Health and Social Welfare has recently issued an accountability tool called the Reproductive, Maternal, Newborn, and Child Health

(RMNCH) Scorecard.[217] The Tanzanian president launched the RMNCH scorecard in May 2014 and tasked the regional (provincial) commissioners to ensure 'greater progress is made in the next 500 days to reduce preventable maternal, newborn and child deaths'.[217] The scorecard has been designed to use HMIS data from districts where reports will be generated on a quarterly basis to inform the progress made on reaching the targets on maternal, newborn and child survival. The scorecard will report on 13 indicators within categories of pre-pregnancy, pregnancy, labour and delivery, newborn health, child health, stock-out data for tracer drugs, data completeness and timeliness, and on the community health fund. It is hoped that leaders in districts/regions will be accountable for the achievement of the objectives. While the RMNCH scorecard is an essential accountability tool with much needed political support, relying on HMIS data for nearly all of the scorecard indicators may not reveal the extent to which districts are progressing [or not]. Monitoring of not only the scorecard indicators but also implementation strength indicators, contextual factors and equity indicators will help highlight areas requiring quick attention, thereby helping regional and district authorities to change their course of action appropriately.

10.3 Overall strength and coverage in districts

Table 10.2 summarises the composite scores of study districts on the implementation strength components and on the coverage of FANC and EmOC programmes including scores of the contextual factors that were significantly associated ($P<0.05$) with the overall implementation strength. The table is focused on the weighted results from aggregated indicators of the six building blocks as active components of FANC and EmOC programmes (implementation strength) and selected proxy indicators of programme coverage and contextual factors. Preference weights generated from the opinions of the maternal health experts on the six building blocks (Figure 3.3 on FANC and Figure 3.7 on EmOC and respective discussion in Section 3.6) strongly aligned with overall contributions of each block as showed by programme data for the implementation strength (Table 5.2 on FANC and Table 5.3 on EmOC). Comparing the corresponding columns (Table 10.2) on implementation strength and programme coverage for FANC and EmOC programmes respectively, results show that coverage scores were higher than implementation strength scores

Table 10.2: Overall Summary of Results

Study district		Overall Implementation Strength		Overall Programme Coverage		Contextual Factors (P<0.05 on implementation strength)				
		FANC	EmOC	FANC	EmOC	Total Fertility Rate	Female literacy level (adult, primary and secondary education)	Households using drinking water from improved water sources	Households using improved toilet facilities	Households affected by famine
Urban Districts	Arusha Urban	86%	80%	133%	92%	4.3	27%	76%	71%	1%
	Ilala	56%	53%	147%	97%	3.2	24%	88%	83%	0%
	Iringa Urban	61%	57%	80%	102%	3.9	21%	94%	80%	1%
	Kinondoni	43%	39%	70%	79%	3.2	18%	73%	55%	0%
	Mtwara Urban	52%	49%	64%	79%	4.8	17%	55%	49%	2%
	Songea Urban	49%	56%	118%	103%	5.0	25%	73%	96%	0%
	Tanga Urban	46%	47%	76%	81%	4.2	21%	87%	81%	0%
	Temeke	51%	42%	109%	63%	3.2	23%	68%	53%	6%
Rural Districts	Babati	38%	40%	61%	62%	4.3	21%	60%	30%	25%
	Bagamoyo	54%	48%	69%	53%	5.0	22%	68%	41%	23%
	Geita	38%	30%	82%	43%	5.7	27%	57%	45%	4%
	Kahama	31%	25%	75%	53%	7.4	19%	61%	47%	17%
	Kasulu	26%	25%	70%	58%	7.5	21%	74%	66%	0%
	Kilosa	33%	39%	56%	53%	5.0	19%	77%	79%	40%
	Kondoa	29%	31%	58%	52%	5.4	23%	33%	56%	38%
	Mbozi	28%	28%	80%	45%	6.2	24%	39%	32%	6%
	Moshi Rural	44%	47%	42%	47%	3.1	21%	74%	31%	5%
	Muleba	29%	32%	82%	72%	7.4	22%	63%	52%	73%
	Musoma Rural	24%	22%	61%	31%	7.0	20%	41%	13%	77%
	Ruangwa	37%	42%	66%	56%	4.0	17%	36%	65%	0%
	Singida Rural	41%	33%	56%	38%	4.9	22%	32%	40%	35%
	Sumbawanga R	27%	28%	83%	56%	7.0	22%	49%	42%	0%
	Uyui	30%	25%	99%	47%	7.3	22%	29%	28%	0%
Overall Mean		41%	40%	80%	64%	5.2	22%	61%	54%	15%
Std Deviation		14%	14%	26%	21%	1.5	3%	19%	21%	23%

across all study districts suggesting that more efforts on implementing the programmes would result in even more coverage of the programmes. Future research would provide better evidence on this result by collecting similar data over time including more indicators on both implementation strength and programme coverage, which was not possible in this study due to resource and time limitations. Calculations and the discussion involving the scores and ranking of the study districts are described in respective chapters (Chapters 4 and 5 on implementation strength, and Chapter 6 on programme coverage).

The linear associations between the overall implementation strength of FANC and EmOC and the five contextual factors (last five columns of Table 10.2) as were noted in Chapter 8 are worth highlighting. FANC implementation strength was significantly associated with the contextual factors of female literacy level and households using improved toilet facilities (Table 8.1), while EmOC implementation strength was significantly associated with the contextual factors of total fertility rate, households using drinking water from improved water sources, households using improved toilet facilities, and households affected by famine (Table 8.2, discussions on Section 8.4). These results suggest that, as confounding factors, the contextual factors could alter the observed association between programme implementation strength and the outcome (programme coverage). So for FANC programme, households with women having higher literacy level or using improved toilet facilities have the potential to change the coverage of FANC programme. Likewise, for EmOC programme, households with women affected by famine, or failure to bring down the TFR, or using drinking water from improved water sources, or using improved toilet facilities may alter the coverage of EmOC programme. As an afterthought, having collected data and presented results on several of these factors, it could be considered that the implementation strength approach demonstrated its capability to capturing and reporting equity dimensions.

While it is well known that female literacy is associated with utilisation of health services,[291] [292] it comes as no surprise that districts with high implementation strength are also associated with low fertility rates, as this outcome is also associated with higher education among women of reproductive age.[293] [294] Given that households with access to better water and sanitation services are contextually an indication of being within a high socioeconomic group,[295] [296] [182] [297] the concern remains only in the relationship between FANC and EmOC implementation

strength and households affected by famine. While these results do not necessarily show direct causal relationships, it is possible that women in households affected by famine are likely to participate in socioeconomic activities more so during periods of famine, potentially affecting antenatal care attendance. It therefore also possible to expect FANC and EmOC programmes to be weak in those areas affected by famine, partly because they are isolated and poor. Thus, while these contextual factors are likely to affect the coverage of programmes, the factors can as be affected by the programmes.

There were incidences of flooding in the three districts of Dar es Salaam, the main city in Tanzania, in early 2011, but findings of this study did not find an association between service access or use and flooding, possibly not because the floods had no effect, but because the data set would not have picked up such an effect. While this study reported the names of other programmes implemented in study districts, there was no clear metric for use in comparing the districts in quantifiable scores. Nevertheless, most of the districts reported to have several ongoing projects, especially on malaria and HIV/AIDS. Some districts reported having research activities related to maternal health and other community-based programmes. Overall, to be able to assess the contextual factors as potential confounding variables between implementation strength and program coverage, it is imperative for large-scale maternal health programmes to collect their data on a regular basis so that changes over time can be detected and included in the analyses.

10.4 Strengthening FANC and EmOC programmes

Results from maternal health experts (Chapter 3) and data from the implementation strength components (Chapter 4), suggest that all six components are important in contributing to the strength of FANC and EmOC programmes. Results from the illustrative dose-response of the overall implementation strength and overall programme coverage (Figures 8.1 and 8.2) also indicated that by scaling up the six blocks, it is possible to increase the coverage of the two programmes. Scaling the overall implementation strength should therefore consider a careful combination of the efforts provided to each block. For example, to ensure that there is an equitable distribution of health workers (*availability* of doctors, nurses and midwives) in study

districts, the national policy needs well-coordinated efforts in producing sufficient annual numbers of the cadres to match with geographical needs. *Competence* of the health workers should be maintained through regular short course training and *motivation* of the workers should be addressed by national and local health authorities. While the number of health facilities is encouragingly sufficient in study districts (as seen in this study, Table 4.1 with discussion on Section 4.4), distribution of the facilities needs to take consideration of the hard-to-reach areas where services should also be made reachable and accessible. In addition, all hospitals and health centres should be equipped with operating theatres to be able to offer comprehensive EmOC services.[203]

While in Tanzania it is the responsibility of health facilities in preparing correct orders of the essential drugs and supplies and according to their local needs, the Medical Stores Department (being the central procurer of medical products)[298] needs to improve its logistics in order to mitigate stock out of the medical products.[164] When there is notable failure (by means of logistics and timeliness) on the Medical Stores Department, health facilities should be allowed to make their purchases of the medical products for future reimbursement. It should be noted that, even though facilities are allocated with budgets for medical products, they do not have outright local control of the monies as this is done centrally by the Medical Stores Department.[299]

Strengthening the quality of FANC and EmOC services might be partly achieved by promoting use of evidence-based guidelines and standards such as the Jhpiego's Standards-Based Management and Recognition Approach (SBM-R).[300] SBM-R involves training of the service providers, provision of supportive supervision, assessments of service quality, and facility-based action plans. Such quality improvement approaches and tools are essential in identifying and solving local constraints in study districts and are more likely to improve the delivery of FANC and EmOC services. While this study did not assess the pattern and levels of Thaddeus and Maine's three delays in districts, coverage of FANC and EmOC programmes could also be linked to the three. The three delays of care (discussed in Section 1.6) are: delay in decision to seek care, delay in getting to health facilities, and delay in receiving care.[130] Strengthening of FANC and EmOC programmes should therefore ensure that community-based programmes are adequately addressing the three delays and are reaching both the women (as primary target

populations) and the men (as key partners involved in household decisions and in providing support). Community-based programmes for maternal health at the national-level and local-level can involve mobilising communities through different channels including advocacy and awareness, social marketing, capacity strengthening, and behaviour change communication.[301]

10.5 The role of partnerships in implementing programmes

Having documented and observed a long list of projects, programmes and initiatives operating in the study districts, some of which have local-, national-, or global-level presence (Table 7.6), it is evident that implementation of FANC and EmOC in the study districts is affected in one way or another by such efforts. For example, projects involving immunisation, micronutrients, nutrition, and control of malaria and HIV/AIDS are likely to influence both the implementation and coverage of FANC and EmOC programmes.[302] [136] It is imperative to ensure that such efforts are well-coordinated. Partnership in maternal health services is diverse in nature but is likely to involve implementation strength and programme coverage, technical partnership and political partnership. Such partnership in Tanzania is likely to include the Ministry of Health and Social Welfare, the Ministry of Finance and Economic Affairs, the Ministry of Regional Administration and Local Governments, donor agencies and maternal health programme implementing partners.

Projects, programmes, and major initiatives operating on the implementation strength side can be those providing health services–related factors such as building health facilities, training community health care workers, construction of shallow wells in communities and provision of drugs and supplies. Partnership on the programme coverage side involves all projects, programmes and major initiatives addressing issues within the three delays and are most likely interlinked with those on the implementation strength side, such as those forming the Safe Motherhood Working Group discussed in Section 3.2.2. Technical partnership (such as the Monitoring and Evaluation Technical Working Group) should be centrally coordinated by the Ministry of Health and Social Welfare in order to effectively and continuously convene technical minds in discussing current evidence-based practices and in analysing and deliberating priorities and efficient approaches for

implementing essential maternal health interventions.[303] Involvement of maternal health professional associations (such as the Tanzania Nursing and Midwifery Council,[304] Tanzania National Nursing Association,[305] the Medical Council of Tanganyika[306] and the Medical Association of Tanzania[307]) and non-governmental organisations is equally important in ensuring balanced participation in technical assistance in maternal health.

Political partnerships on the other hand, should involve convening multi-stakeholder dialogue processes that involve issues on policy, effective implementation and promotion of maternal health. Such efforts should not only involve high profile and national-level individuals and organisations (primarily those enabling policies, with political will, and with strong capacity for advocacy) but also it should encourage participation of low-level community groups involved in maternal health promotion. Furthermore, because most African health ministries have limited human resources to fully monitor and coordinate all efforts on women's health involving a multitude of players,[308] its central role should therefore be on setting priorities and implementation strategies for all donors and implementing partners to adhere to. Likewise, the Tanzanian Ministry of Finance and Economic Affairs (being the central funding body for the annual comprehensive council (district) health plans) has a vital role to play through working with the Ministry of Health and Social Welfare to balance and prioritize the financing of maternal health initiatives.

In addition, the Ministry of Regional Administration and Local Governments (under which hiring of local staff, local health planning and implementation is managed) should also be actively involved with the ministries of finance and health in ensuring that information on cost-effectiveness of the maternal health interventions is shared to ensure projects are acceptable, equitable and likely to be feasibly implemented.[309] Although strong partnership is essential in all of the above described factors for effective and successful implementation of maternal health programmes, sustained financial support, strong communication among implementing partners and local participants and continuous monitoring and evaluation of the programmes is equally important for improvement in policy issues and high coverage of the programmes.[310] [311]

10.6 Bias, limitations and methodological challenges

There were several identified sources of bias for this study, including the selection process of the maternal health experts, the instruments used and the proposed duration of FANC and EmOC programmes. In addition, questionnaires used as research instruments in collecting data were also likely to introduce some level of measurement bias in the event that they did not ask the 'right' questions or prompt appropriate responses, thereby not being able to detect differences in the variables of interest. Sufficient time-series data for both the implementation strength, programme results and the contextual factors was another limitation that did actually affect the results of this study.

In an effort to reduce bias, data were collected from different types of maternal health experts. By so doing, the study avoided having perspectives from only one cadre within the health system. To improve the validity and reliability of using expert judgment, the study used actual programme data it collected from study districts for cross-validation. A similar approach in cross-validating expert opinions was reported in a study conducted in Kenya and Bangladesh.[312] The study used different methodologies such as print media, radio programmes and interviewing representatives from different ministries. Results of these studies were in some ways similar to those originating from the experts in this study. Even though health facilities are important contact points for users of FANC and EmOC services, it is always difficult to get accurate denominator populations for service utilization, coverage and impact. It is therefore important to use data from population surveys to leverage indicator reporting. For example the issue of facility coverage scores exceeding 100% (noted in several parts of the results chapters) could alternatively use denominators that are best estimated through a complete vital registration system. Nevertheless, some facility-based indicators, such as the percentage of institutional deliveries, tend to give better population characteristics, as the number of annual expected live births is used as the population-based denominator compared to other indicators such as the indicator on the rate of HIV/AIDS testing, which uses only the number of first antenatal care visits for the denominator. The study used the number of first antenatal care visits to assess the coverage of FANC programmes instead of four or more antenatal care visits. Even though both numerators are

recommended by the WHO, a pregnant woman is considered fully protected and prepared for safe delivery if she completes four or more antenatal care visits. By not using the latter as a numerator on FANC coverage, it is possible that some study districts would have scored higher and some lower than their current first visit-based scores.

In the theory of evaluation studies, resources, inputs and processes are expected to produce the intended results: immediate (outputs), intermediate (outcomes) and eventually long-term (impacts). However, besides implementation of programme activities, there are a myriad of conditions that need to be met in order for the intended results to be realised or achieved. Examples of the resources, inputs and processes required for both FANC and EmOC programmes include, but are not limited to: having an adequate number of health workers who are sufficiently competent, responsive and productive; having well-stocked essential drugs and supplies for antenatal care and emergency obstetric care throughout the year, including an improved integrated logistics system for timely delivery; districts are allocated sufficient funding from the central government that is made available on time to facilitate programme execution; a sufficient number of health facilities provide antenatal and emergency obstetric care and are geographically well distributed and easily accessible; and an adequate number of health facilities have the capacity to manage all the nine signal functions. Use of coverage indicators for programme results can sometimes obscure the actual causal relationship between the implementation strength and programme results, where one coverage indicator may correlate with another coverage indicator. Care should thus be taken in identifying such correlations and account for them during data analysis.

Although the study collected data on district revenues and expenditure, there were no costing data on programme resources and other inputs as it was beyond the scope of this thesis. Collection of costing data could have been of great use in evaluating the cost-effectiveness of the programmes. It is possible that some policymakers and donors associate the impact/effect of programmes with cost. In addition, further thought is also required in developing a composite index for other programmes operating in study districts to be able to assess their influence over the implementation of maternal health in districts. Other than documenting the list of such programmes operating in study districts, the study did not quantify their effect in the form of a composite index or estimate their contribution in the implementation

strength of the two tracer programmes. It was previously reported (Section 8.3) that not all indicators for the implementation strength, programme coverage and contextual factors had data for both 2010 and 2011. Lacking data for all indicators and for the entirety of both years denied the use of all possible statistical analyses. Continuous documentation of programme activities, contextual factors, and programme results is recommended for at least three years to be able to generate more significant results of the associations between programme efforts and programme results.

10.7 Conclusion and the way forward

This final section summarises the methodological and pragmatic contributions made by this thesis including the final thoughts on potential further research. There were two main methodological contributions of the work undertaken in this thesis: The first was the development and testing of the implementation strength approach for use in evaluating large-scale maternal health programmes in low- and middle-income countries. The second was testing the approach's validity regarding the contents of its tools and its overall acceptability for use by national coordinators and district health authorities. This thesis introduced an alternative method for evaluating large-scale maternal health programmes that has the potential for application in low- and middle-income countries. At the planning stage, programme evaluators need to identify all possible components that are likely to have the greatest impact within maternal health programmes. They also need to prepare a list of proxy indicators for the programme components (programme implementation strength), for the anticipated programme results, and for the contextual factors. The indicators will guide data collection and programme monitoring. At the programme's end, evaluators will need to conduct numerous analyses for all data collected, including dose-response to detect the association between the implementation strength and programme results.

Dose-response results can help inform programme implementers and donors on the estimated inputs required to achieve specific levels of results. By knowing the strength of programmes, it is possible to usefully compare results of different programmes, thereby helping donors to realise different funding modalities for maternal health programmes that will yield more health gains for every dollar invested. For programme implementers, measuring the implementation strength

scores might be useful in assessing the coverage and the intensity of maternal health programmes. In addition, knowing the effect of the contextual factors in programme implementation, those responsible for programme execution can gain insights on the amount of resources required to maintain a certain level of programme activities towards the achievement of the intended results. The use of six health system building blocks and testing their contributing strength into the implementation of FANC and EmOC programmes was likely undertaken for the first time in Tanzania. It is hoped that the learning gained from developing the approach and its attempted application to measuring the strength of the two programmes in 23 study districts in Tanzania would be useful for others hoping to use similar evaluation approaches for large-scale maternal health programmes in low- and middle-income countries.

The higher the coverage of FANC and EmOC programmes the better the reliance of vertical programmes using them for delivering interventions.[207] Even though this study found high coverage of the FANC programme in study districts (scoring an overall average of 80%), there was no information on equity distribution of coverage among different socioeconomic groups besides the study showing that urban districts scored much better than rural districts. Inequity in accessing maternal health services has unfavourable effects on users, especially the young, those living in rural settings, the poor and those with no or less education.[26 313 236] Advocating for high coverages of FANC and EmOC programmes should therefore take the view of the universal health coverage and should focus on extending it to those not currently covered by the programmes, and on reducing cost-sharing and user fees as well as including other services.[314] Further research extending this work should therefore ensure it captures equity dimensions among its indicators to be able to describe the effect of maternal health programmes in all socioeconomic groups in populations.

There is also a need for further research in identifying the key components of a large-scale maternal health programme that will inform the most effective scale-up implementation approaches as well as how to best monitor and evaluate targeted high-impact interventions across the continuum of care. For maternal health, high impact interventions for Tanzania and for most of the 75 high-burden priority countries include use of modern contraceptives (during the pre-pregnancy period), more antenatal care visits to clinics (during the pregnancy period), delivering in health facilities where there is quality labour and delivery management (during childbirth), and postpartum and postnatal care within two days of delivery (after

childbirth). In order to address them adequately, scale-up issues of coverage, equity and quality of essential care need to be rigorously identified, monitored, and evaluated. It was argued earlier (Section 3.1) that, due to the multifaceted nature of the health system in which large-scale interventions and programmes are implemented, obtaining evidence of the effect or impact of the programmes can be challenging. This limitation is partly due to the challenge of getting reliable data from programme activities, which dissuades effective tracking of implementation or scaling-up of programmes and/or measurement of their impact. It was also argued that the essential part of tracking a programme is to identify and continuously document and measure key areas or components that are likely to cause the most impact. While tools such as the Tanzanian RMNCH scorecard or the LiST can be used to evaluate current achievements, further research of the implementation strength approach is recommended. This research can be done by testing the hypothesis that the scaling of targeted high-impact interventions is independent of factors in the context of maternal health programme activities.

In addition to monitoring key programme activities, research on the effectiveness of the approach can be conducted through assessing contextual factors that will provide evidence for any associations between maternal health interventions and their effects. The literature on studies that have investigated and evaluated contextual factors in relation to scaling-up implementation of maternal health programmes in low- and middle-income countries is limited. As it is in the interest of donors and governments to scale up maternal health programmes effectively and cost-efficiently, it is important to investigate, document and analyse the levels of factors within a specific context. Such understanding may better be revealed through research on the approach. If contextual factors are not properly documented, it may not be possible to estimate the effect of specific programmes on maternal health outcomes at scale. One other area that will require further research is on how to generate an index to reflect the contribution of other programmes operating in districts on the implementation strength of FANC and EmOC programmes. The research work needed could involve creating the index as function of the number, duration of operation and coverage scope of the programmes operating in study districts that will be used to estimate their contribution attributable to the strength of the two tracer programmes. Even more research can be conducted to look at how such programmes external to FANC and EmOC programmes are likely to facilitate programme coverage in target populations.

Appendices

Appendix 1: Maternal Health Experts Tool (Implentation Strength Instrument)

INFORMATION SHEET FOR MATERNAL HEALTH EXPERTS

STUDY TITLE: "Towards a new method for evaluating national maternal health programmes in Tanzania: measuring implementation strength of focused antenatal care and emergency obstetric care"

(To be read by the Interviewee)

Why am I conducting this research?

My name is Gregory Kabadi studying at the London School of Hygiene and Tropical Medicine also working with the Ifakara Health Institute in Tanzania. I am interested in developing and testing an approach for estimating programme implementation strength of Focused Antenatal Care (FANC) and Emergency Obstetric Care (EmOC) programmes in the context of the Tanzanian health system that will help to understand the effects these programmes have in targeted populations. Understanding and measuring programme implementation strength may disclose why some programmes have an impact on coverage and health outcomes while others do not. There are no commonly standardized methodologies for measuring programme implementation strength as most approaches are tailored to individual programmes. Programme implementation strength is also known as programme intensity – which refers to programme dose (how much), specificity (of which specific activity/element) and duration (for how long) a programme is implemented.

Measuring programme strength requires an understanding of how programmes work and involves defining measurable concepts, identifying sources of programme data and close follow up of programme activities. I intend to have about three meetings with you and other maternal health experts to identify FANC &EmOC programme elements you consider most important using a framework of the World Health Organization's health system building blocks. I will also conduct interviews with some district-level participants to collect relevant programme data from a nationally representative sample of 23 districts of Tanzania. I will use FANC and EmOC programme data to cross-validate the information gained from maternal health experts.

Study Investigator:

Mr Gregory Kabadi, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT. Tel +255 784 444668; email: Gregory.Kabadi@lshtm.ac.uk

CONSENT FORM FOR MATERNAL HEALTH EXPERTS

How long will the interviews take?

The two anticipated interviews will be conducted in joint sessions and should take approximately 20 – 40 minutes.

Are there advantages or disadvantages involved in taking part?

There are no individual benefits to taking part, but in participating to the joint meetings of maternal health experts in Tanzania, you will help me improve the approach for measuring the strength of FANC and EmOC programmes in our country – that might be of help to national programme managers involved in planning and managing such programmes for efficiency and effectiveness.

Who will have access to the information I give?

The information you give will be for research purposes only. I will not disclose your identity, or use your name in any reports/publications of this work.

What will happen if I refuse to participate?

Participation to this study is voluntary. You are free to decide if you want to take part or not. If you do agree, you can change your mind at any time. You can refuse to answer any specific questions, or stop the interview at any time. If you choose not to answer a question, stop the interview, or even not participate at all in the study, there will not be any adverse consequences for you or your organisation.

What if I have any questions?

If you have any questions, you can ask them during the joint meeting. If you wish to ask questions later, you may contact me by email at:

Study Investigator: Mr Gregory Kabadi, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT. Tel +255 784 909067; email: Gregory.Kabadi@lshtm.ac.uk

Sharing the results

Findings from doing this study will be shared with you through another joint meeting and/or in any relevant national-level maternal health meetings. After analysis, I shall publish the results in order that other interested people may learn from our experience.

Interviewee's Consent

I have read the above information. I have been able to ask questions about it and any questions that I have asked have been answered to my satisfaction. I consent voluntarily to participate in this study.

Signature of the Respondent: _____

Respondent's Designation:_____

Date of signature for consent to participate ___/___/2012

Weighting & Ranking Tool for Maternal Health Experts

Instructions on how to weigh/rank FANC &EmOC Programmes

For each of the two FANC and EmOC programmes, there are two sections to fill in information regarding what you think contributes to the strength of programmes. Section 1 is about programme elements for FANC and Section 2 is about programme elements for EmOC:

1. There are 6 categories for each programme (FANC or EmOC). On a scale between 0% and 100%, please estimate the weight of each category in a manner that you think it contributes to the strength of a programme (FANC or EmOC). **All category weights should add up to 100%**.

2. For each of the 6 categories of a programme, there are elements/activities that constitute the category. Please rank these elements within a category in the order of their importance/priority. For example, if there 5 elements in a category, then elements should rank from 1st to 5th.

A: Focused Antenatal Care

No	Category	Category Weight (%)	Programme Elements	Element Rank
1	Human Resource		**1. Availability**: a) Enrolment of pre-service nurse-midwife students in allied health colleges b) Establishing and deploying Community Health Workers for Maternal care	
			2. Competence: c) In-service training of service providers on FANC d) Empowering TBAs through capacity building in disadvantaged areas	
			3. Responsiveness & Productivity: Workplace motivation/incentives for HRH e.g. staff housing, hardship allowance, timely salaries and staff loans, etc.	
2	Drugs and Supplies (& Logistics)		1. Procurement of Essential (Tracer) Drugs for Antenatal Care	
			2. Procurement of Essential Equipment & Supplies for ANC	
			3. Improving Integrated Logistics System (ILS) for drug supplies	
3	Service Provision (Incl. Infrastructure)		1. Improving ANC service provision (including PMTCT & FP)	
			2. Referrals – within and to higher health facilities.	
			3. Building new RCH Clinics; Renovating/regular maintenance of old/existing buildings	
			4. Providing Incentive Vouchers (e.g. a pair of Kanga) to attract women to deliver in health facilities	
			5. Behavioural change communication programmes	
			6. Community male involvement	
			7. Community empowerment (e.g. accountability of service providers/CHMTs to community complaints, etc.)	
4	Health Financing		1. Increasing percentage of the Gov health budget spent on: • Salaries of health workers • Medicines and supplies • Other recurrent costs (e.g. admin cost at MOHSW, CHMT, In-service training)	
			2. Increasing Health Insurance (Coverage, Funding, Policy) • Social Health Insurance – NHIF & CHF • Private for profit	
5	Health Information System		**1. Availability:** Timely reporting of FANC information through quarterly reports from health facilities to the district	
			2. Use: Incorporating routine information and evidence from research into district (and/or facility) for planning purposes on FANC	
			3. Capacity/Capability: • Purchasing and installing ICT equipment for data entry and data management in facilities/districts • Adequate # of registers & reporting forms in facilities	
6	Leadership/ Governance		1. National Level advocacy for maternal health at all levels	
			2. Supportive supervision visits to health facilities	
			3. Issuance of guidelines, rules and regulations – licensure of health professionals, accreditation & certification	
			4. Development by Decentralization (DbyD)/Basket Fund use Vs Local Gov funding policy	
			5. Involving health and non-health stakeholders in defining and prioritizing health needs and services at the national level	
7	Other (Please specify...)		1.	
			2.	
			3.	
			4.	
			5.	
			6.	

B: Emergency Obstetric Care

No	Category	Category Weight (%)	Programme Elements	Element Rank
1	Human Resource		**1. Availability**: a) Enrolment of pre-service nurse-midwife students in allied health colleges b) Allocation of sufficient number of Anaesthetists, Laboratory Technicians, Nurse-Midwives and Clinical Officers in Comprehensive EmOC facilities c) Establishing Community Health Workers for Maternal care and Life SS	
			2. Competence: a) In-service training of service providers on EmOC/IMPACC b) Practical training exposure to Clinical Officers and AMOs on EmOC in hospitals with Specialists; Continuing Medical Education and On-the-job-training for all c) Empowering TBAs through capacity building in remote areas	
			3. Responsiveness & Productivity: Workplace incentives for HRH e.g. staff housing, hardship allowance, timely salaries and staff loans, monetary/non-monetary motivation, etc.	
2	Drugs and Supplies		1. Procurement of Essential EmOC Drugs	
			2. Procurement (and maintenance) of Essential Equipment & Supplies for EmOC	
			3. Availability of emergency transport for patients; and for workers during emergency	
			4. Improved Integrated Logistics System (ILS) for drug supplies	
3	Service Provision (Incl. Infrastructure)		1. Upgrading the existing health centres to enable them to provide both Basic and Comprehensive EmOC signal functions (building Operating Theatres, L&D wards)	
			2. Establishment of maternity waiting homes in hospitals	
			3. Allocating providers with houses w/in facility grounds for quick emergency calls	
			4. Radio call communications & ambulances in selected health facilities	
			5. Ongoing BCC programmes & community mobilization/empowerment initiatives	
			6. Outreach/Mobile services	
			7. Availability of Safe Blood	
			8. Staff meetings	
			9. Waste safe disposal	
			10. Water and Power reliability (source, storage & emergency)	
4	Health Financing		1. Increased government health budget for expenditure on: • Salaries of health workers • Medicines and supplies • Other recurrent costs (e.g. admin cost at MOHSW, CHMT, In-service training)	
			2. Increasing Health Insurance (Coverage, Funding, Policy) • Social Health Insurance – NHIF; CHF; Private for profit	
5	Health Information System		**1. Availability:** • Timely reporting of EmOC information through quarterly reports • Conducting maternal death audits	
			2. Use: Incorporating routine information and evidence from research (and maternal death audits) into district (and/or facility) for planning purposes on EmOC	
			3. Capacity/Capability: • Purchasing and installing ICT equipment for data entry and data management in facilities/districts • Adequate number of registers and reporting forms in facilities	
6	Leadership/ Governance		1. National Level advocacy for maternal health at all levels	
			2. Supportive supervision visits to health facilities	
			3. Issuance of Guidelines, Charts and Algorithms (Wall Charts), rules & regulations	
			4. Involvement of health and non-health stakeholders in defining and prioritizing health needs and services at the national level	
			5. Development by Decentralization (DbyD)/Basket Fund use Vs Local Gov funding	
7	Other (Please specify…)		1.	
			2.	
			3.	

Appendix 2: Data Collection Form for Contextual Factors in a District

1. Health

(Main source of data is the District Medical Officer's Office - including the District Health Secretary's Office AND/OR the District Social Welfare Department)

No	Project/Programme type	List of projects/programmes in the district	Main purpose	Area of service	Main financial provider
1	Other maternal and newborn health programmes	e.g. MAISHA programme	Maternal & Newborn health	23 facilities (of 56)	Jhpiego
		1)			
		2)			
		3)			
		4)			
2	Child health programmes (immunization, micronutrients, nutrition,	e.g. MWANZO BORA programme	Reducing the prevalence of stunting in under 5	4 villages in the district	Africare
		1)			
		2)			
		3)			
3	Malaria projects/programmes (indoor residual spraying, bed net use, etc.)	e.g. COMMIT project	Behaviour Change Communication for malaria	10 wards	JHU-CCP
		1)			
		2)			
		3)			
		4)			
		5)			
4	HIV/AIDS projects/programmes (Care and Treatment, Counselling and Testing, etc.)	e.g. ANGAZA/ANGAZA ZAIDI	Voluntary Counselling and Testing	26 village centres	AMREF
		1)			
		2)			
		3)			
		4)			
		5)			
		6)			
		7)			
5	Community health/Social protection programmes	e.g. PAMOJA TUWALEE	Social protection for Most Vulnerable Children	4 wards (of 11 wards)	PACT/TASAF
		1)			
		2)			
		3)			
		4)			

2. Water

What percentage (actual or likely) of the population/households used drinking water from improved water sources in the district?

2010: |___|% **2011:** |___|%

- Definition/Clarification: Improved drinking water sources include: Piped water into dwelling, Piped water to yard/plot, Public tap or standpipe, Tube well or borehole, Protected dug well, Protected spring and Bottled water)
- Data Source/s: District's Planning Office; Water Department

3. Sanitation

What percentage (actual or likely) of the population/households used improved toilet facilities in the district?

2010: |___|% **2011:** |___|%

- Definition/Clarification: Improved toilet facilities include: A flush toilet, A piped sewer system, A septic tank and A protected or covered or ventilated pit latrine
- Data Source/s: District's Planning Office; Sanitation & Environment Department

4. Transport

a) What percentage (actual or likely) of villages/streets with accessibility to public (or own) transport services within 3 kilometres of district road's network?

2010: |___|% **2011:** |___|%

Tranport...

b) What percentage of villages/streets whose roads were impassable during rainy season?

2010: |_16__|% **2011:** |5___|%

- Definition/Clarification: Especially transportation and proximity to reliable health facilities
- Data Source/s: District's Planning Office; Transportation Department

5. Communication:

What percentage (actual or likely) of villages/streets was covered by the available telecomm services and networks in the district?

2010: |_86__|% **2011:** |87___|%

- Definition/Clarification: Available statistics (or estimates) on use of telecom services and networks such as TTCL, Airtel, Voda, Tigo, etc.
- Data Source/s: District's Planning Office; Communication Department

6. Natural Disasters:

What percentage (actual or likely) of households was affected by natural disasters?

No.	Type of disaster	2010 (%)	2011 (%)
1	Drought		
2	Floods		
3	Disease outbreaks		
4	Famine		

- Definition/Clarification: Any available data estimating the percentage of households affected by natural disasters in the district
- Data Source/s: District Medical Officer; Department of Agriculture

Appendix 3: Indicators for Data Extraction from DHIS Database

District Name: _____

	No	Data Item	2010	2011
Antenatal Care	1	Number of first ANC visits		
	2	Number of ANC clients with FOUR visits		
	3	Number of ANC clients with 2-5 tetanus toxoid injection doses		
	4	Number of ANC clients receiving 2 DOSES of IPT (SP)		
	5	Number of ANC HIV tests done		
	6	Number of ANC clients with positive HIV tests		
	7	Number of ANC clients receiving 90 tablets of IRON (Fe) or FeFol		
	8	Number of ANC clients receiving Mebendazole/Albendazole tablets		
	9	Number of ANC clients who tested for Haemoglobin		
	10	Number of ANC clients screened for Syphilis		
	11	Number of ANC clients who received ITN vouchers		

		Number of DAYS of STOCK OUT **per quarter**:	2010				2011			
			J-M	A-J	J-S	O-D	J-M	A-J	J-S	O-D
	12	Tetanus Toxoid								
		SP Tablets								
		Fe/+Fol Tablets								
		Mebe/Albendazole								

	No		2010	2011
Obstetric Care	1	Number of women giving birth at this facility		
	2	Number of women with major direct obstetric complications treated		
	3	Number of **Caesarean sections** performed		
	4	Number of maternal deaths due to **direct** obstetric causes		
	5	Number of maternal deaths due to **indirect** causes		
	6	Number of intrapartum deaths		
	7	**Within the last 3 months**, has there been any reported cases of the following signal functions of complicated delivery **in this facility**:	√ = YES x = NO	√ = YES x = NO
		Parenteral administration of antibiotics		
		Administration of uterotonic drugs		
		Administration of parenteral anticonvulsants for pre-eclampsia & eclampsia		
		Manual removal of the placenta		
		Removal of retained products		
		Assisted vaginal delivery		
		Basic neonatal resuscitation		
		Caesarean section		
		Blood transfusion		

		Number of DAYS of STOCK OUT **per quarter**:	2010				2011			
			J-M	A-J	J-S	O-D	J-M	A-J	J-S	O-D
	8	Oxytocins								
		Ergometrine								
		Magnesium Sulphate								

Appendix 4: Assessing Content &Face Validity of the Tool for Measuring Implementation Strength

A. Briefing:

The assessment tool has been designed as a potential aid for district health managers and planners. It is intended to help them identify components of maternal health projects that have made a greater contribution to strength to implementation. Please spare a few minutes to answer the following questions concerning the assessment tool. This should take you no more than five minutes.

B. About you:

1. Job Title: _____

2. Total work experience: ☐ Under 5 yrs ☐ 6 – 15 yrs ☐ 16 – 25 yrs ☐ 26 yrs or more

3. Place of work: (Tick all that apply) ☐ Public ☐ NGO ☐ Private ☐ Other

4. Education: ☐ Certificate ☐ Diploma ☐ Adv ☐ Masters or

5. Age: ☐ Under ☐ 25 - 40yrs ☐ 41 – 55yrs ☐ 56yrs or

6. Sex: ☐ Female ☐ Male

C. Your opinion of the tool:

Question	Response						
For each item, please tick the circle representing your response	*Strongly Disagree*						***Strongly Agree***
A. Algorithm:							
i. The **overall** algorithm of the tool is easy to understand and follow	1	2	3	4	5	6	7
B. FANC measurements:							
i. Each component has just about the right elements	1	2	3	4	5	6	7
ii. The algorithm is easy to understand and follow	1	2	3	4	5	6	7
iii. There is an element missing, *name:* _____of which component _____							
C. EmOC measurements:							
i. Each component has just about the right elements	1	2	3	4	5	6	7
ii. The algorithm is easy to understand and follow	1	2	3	4	5	6	7
iii. There is an element missing, *name:* _____of which component _____							
D. Overall tool assessment:							
i. The tool is suitable for use in assessing strength of implementation	1	2	3	4	5	6	7
ii. The instructions are easy to follow (understandable)	1	2	3	4	5	6	7
iii. The length of the tool is reasonable	1	2	3	4	5	6	7
iv. The scoring system and interpretation of the scores is clear	1	2	3	4	5	6	7
v. I would use the tool myself	1	2	3	4	5	6	7
E. General comments: (Do you have any suggestions that will help improve the tool?)							

Bibliography

Reference

1 Dr Margaret Chan, Director-General of the World Health Organization. Opening remarks at the Fourth Global Meeting of Heads of WHO Country Offices, Geneva, Switzerland. 2007. Available here: http://www.who.int/dg/speeches/2007/20071112_geneva/en

2 World Health Organisation. Key facts on maternal mortality (Internet access, 31th Oct 2014), at: http://www.who.int/mediacentre/factsheets/fs348/en/

3 United Nations Millennium Development Goals. Accessed at http://www.un.org/millenniumgoals/bkgd.shtml on 31st Oct 2014.

4 Trends in maternal mortality: 1990 to 2013. Estimates by WHO, UNICEF, UNFPA, the World Bank and the United Nations Population Division. The World Health Organisation, 2014. Geneva

5 Institute for Health Metrics and Evaluation (IHME). Financing Global Health 2012: The End of the Golden Age? 2012. (Institute for Health Metrics and Evaluation). http://www.healthmetricsandevaluation.org/publications/policyreport/financing-global-health-2012-end-golden-age

6 Borghi J, Ensor T, Somanathan A, Lissner C, Mills A, on behalf of The Lancet Maternal Survival Series steering Group. Maternal Survival 4 Mobilising financial resources for maternal health. Lancet 2006; 368: 1457–65 Published Online September 28, 2006 DOI:10.1016/S0140-6736(06)69383-5

7 Costello A, Osrin D. The case for a new Global Fund for maternal, neonatal, and child survival. Lancet 2005; 366:603–05.

8 Ensor T, Ronoh J. Effective financing of maternal health services: a review of the literature. Health Policy 2005; 75:49–58.

9 Lawn JE, Cousens S, Bhutta ZA, et al. Why are 4 million newborn babies dying each year? Lancet 2004; 364:399–401

10 Graham WJ. The global problem of maternal mortality: inequalities and inequities. In: MacLean AB, Neilson J, eds. Maternal morbidity and mortality. London: Royal College of Obstetricians and Gynaecologists, 2002: 3–20

11 Kunst AE, Houweling T. Inequalities in the utilisation of delivery care and other maternal health services. Studies Health Serv Organ Policy 2001; 17:297–315.

12 Fraser A, Kamal-Smith M, Watkins K, The cost of childbirth: how women are paying the price for broken promises on aid. Oxfam briefing paper 52. Oxford: Oxfam, 2004

13 Graham WJ. Now or never: the case for measuring maternal mortality. Lancet 2002; 359:701–04.

14 Parkhurst JO, Danischevski K, Balabanova D. International maternal health indicators and middle-income countries: Russia. BMJ 2005; 331:510–13.

15 Borghi J, Ensor T, Neupane B, Tiwari S. Financial implications of skilled attendance at delivery: the case of Nepal. Trop Med Int Health 2006; 11:228–37

16 Stenberg K, Axelson H, Sheehan P, Anderson I, Gulmezoglu AM, Temmerman M, et al, on behalf of the Study Group for the Global Investment Framework for Women's Children's Health. Advancing social and economic development by investing in women's and children's health: a new Global Investment Framework. The Lancet. 2013. Published online: doi:10.1016/S0140-6736(13)62231-X

17 The World Bank Group. Press Release Issued on September 23, 2013. Accessed on 27th Jan 2014 at http://www.worldbank.org/en/news/press-release/2013/09/23/world-bank-group-invest-700-million-2015-improve-women-children-health-poor-countries

18 World Bank population projection for 2013 for sub-Saharan Africa (developing only). Internet access, on 25th Oct 2014, at http://data.worldbank.org/region/sub-saharan-africa

19 Horton R, Lo S. Investing in health: why, what, and three reflections. The Lancet, Volume 382, Issue 9908, Pages 1859 - 1861, 7 December 2013: doi:10.1016/S0140-6736(13)62330-2.

20 Freedman LP, Graham WJ, Brazier E, Smith JM, Ensor T, Fauveau V, Themmen E, Currie S, Agarwal K. Practical lessons from global safe motherhood initiatives: time for a new focus on implementation. Lancet. 2007 Oct 13;370(9595):1383-91.

21 Filippi V, Ronsmans C, Campbell OMR, Graham WJ, Mills A, Borghi J, Koblinsky M, Osrin D. Maternal health in poor countries: the broader context and a call for action. Lancet, 2006, Volume 368, Issue 9546, Pages 1535–1541

22 World Health Organization, United Nations Children's Fund. Countdown to 2015. Accountability for maternal, newborn and child survival: The 2013 Update.

23 Betrán AP, Merialdi M, Lauer JA, Bing-Shun W, Thomas J, Van Look P, Wagner M. Rates of caesarean section: analysis of global, regional and national estimates. Paediatr Perinat Epidemiol. 2007; 21 (2):98-113.

24 Althabe F, Sosa C, Belizán JM, Gibbons L, Jacquerioz F, Bergel E. Cesarean section rates and maternal and neonatal mortality in low-, medium-, and high-income countries: an ecological study. Birth. 2006; 33 (4): 270-7.

25 Ronsmans C, Holtz S, Stanton C. Socioeconomic differentials in caesarean rates in low- and middle-income countries: a retrospective analysis. Lancet 2006; 368 (9546): 1516–23.

26 Gabrysch S, Campbell OMR. Still too far to walk: Literature review of the determinants of delivery service use. BMC Pregnancy and Childbirth 2009, 9:34 doi:10.1186/1471-2393-9-34

27 Kruk ME, Galea S, Prescott M, Freedman LP. Health care financing and utilization of maternal health services in low- and middle-income countries. Health Policy and Planning. 2007; 22:5, Pp. 303-310.

28 Mills A, Brugha R, Hanson K, McPake B. What can be done about the private health sector in low-income countries? Bull WHO. 2002:80:4. http://dx.doi.org/10.1590/S0042-96862002000400012

29 The World Health Organization. The human resources for health crisis. 2014. [Web material] Accessed 25.04.2014 at: http://www.who.int/workforcealliance/about/hrh_crisis/en/

30 The World Health Organization. The World Health Report, 2006: Working together for health. Geneva, 2006. Report available at: http://www.who.int/whr/2006/whr06_en.pdf?ua=1

31 Mæstad O. Human Resources for Health in Tanzania: Challenges, Policy Options and Knowledge Gaps. The CMI REPORT. 2006. Available online at: http://www.cmi.no/publications/file/2175-human-resources-for-health-in-tanzania-challenges.pdf

32 Kieffer, M. P. and K. Kiragu (2012). "Human resources for health: A key component to achieve the goals of the Global Plan". Paper presented at Global Steering Group Mid-Term Review, IATT Task Force on Human Resources for Health. Nairobi, Kenya. December 6-, 2012.

33 Munga M, Songstad NG, Blystad A, Mæstad O. The decentralisation-centralisation dilemma: Recruitment and distribution of health workers in remote districts of Tanzania. BMC International Health and Human Rights, 2009:9: 9

34 Munga MA, Torsvik G, Mæstad O. Using incentives to attract nurses to remote areas of Tanzania: a contingent valuation study. Health Policy Plan. 2014 Mar;29(2):227-36. doi: 10.1093/heapol/czt007. Epub 2013 Mar 11.

35 Rao M, Rao KD, Kumar AKS, Chatterjee M, Sundararaman T. Human resources for health in India. The Lancet, 2011:377:9765 Pp 587 – 598. doi:10.1016/S0140-6736(10)61888-0

36 Van Lerberghe W, De Brouwere V. Of blind alleys and things that have worked: history's lessons on reducing maternal mortality. Studies Health Serv OrganPolicy 2001; 17: 7–33.

37 Weeks A. Maternal mortality: it's time to get political. BJOG 2007;114:125–126

38 Shiffman J. Generating political priority for maternal mortality reduction in 5 low- and middle-income countries. Am J Public Health. 2007 May;97(5):796-803. Epub 2007 Mar 29.

39 Shiffman J, Okonofua FE, Ved R. The state of political priority for maternal mortality reduction in Nigeria and India. March 1, 2006: http://paa2006.princeton.edu/papers/60071 (accessed Oct 31, 2014).

40 Campbell OMR and Graham WJ. Strategies for reducing maternal mortality: getting on with what works. The Lancet, 2006; 368: 1284–99 DOI:10.1016/S0140-6736(06)69381-1

41 Mutale W, Chintu N, Amoroso C, Awoonor-Williams K, Phillips J, Baynes C, Michel C, Taylor A, Sherr K; Population Health Implementation and Training – Africa Health Initiative Data Collaborative. Improving health information systems for decision making across five sub-Saharan African countries: Implementation strategies from the African Health Initiative. BMC Health Serv Res. 2013;13 Suppl 2:S9. doi: 10.1186/1472-6963-13-S2-S9. Epub 2013 May 31.

42 Boerma JT, Sommerfelt AE. Demographic and health surveys (DHS): contributions and limitations. World Health Stat Q. 1993;46(4):222-6.

43 Merdad L, Hill K, Graham W. Improving the measurement of maternal mortality: the sisterhood method revisited. PLoS One. 2013; 8(4):e59834. Doi: 10.1371/journal.pone.0059834. Epub 2013 Apr 2.

44 Curtis SL, Sutherland EG. Measuring sexual behaviour in the era of HIV/AIDS: the experience of Demographic and Health Surveys and similar enquiries. Sex Transm Infect. 2004 Dec; 80 Suppl 2:ii22-7.

45 Dumont A, Gueye M, Sow A, Diop I, Konate MK, Dambé P, Abrahamowicz M, Fournier P. [Using routine information system data to assess maternal and perinatal care services in Mali and Senegal (QUARITE trial)].[Article in French]. Rev Epidemiol Sante Publique. 2012 Dec;60(6):489-96. doi: 10.1016/j.respe.2012.05.005. Epub 2012 Oct 31.

46 Iguiñiz-Romero R, Palomino N. Data do count! Collection and use of maternal mortality data in Peru, 1990-2005, and improvements since 2005. Reprod Health Matters. 2012 Jun;20(39):174-84. doi: 10.1016/S0968-8080(12)39605-5.

47 Jain AK, Sathar Z, Salim M, Shah ZH. The importance of public sector health facility-level data for monitoring changes in maternal mortality risks among

communities: the case of pakistan. J Biosoc Sci. 2013 Sep;45(5):601-13. doi: 10.1017/S0021932013000126. Epub 2013 Mar 26.

48 Countdown to 2015, Health Metrics Network and World Health Organization. Monitoring maternal, newborn and child health: understanding key progress indicators. World Health Organization, 2011. Report available at: http://www.who.int/healthmetrics/news/monitoring_maternal_newborn_child_health.pdf

49 Unpublished data by McKinsey and Company. Retrieved 10th Jan 2015 from AbouZahr C, Boerma T. Health information systems: the foundations of public health. Bull World Health Org. 2005;83:578-58 at: http://www.who.int/bulletin/volumes/83/8/578.pdf

50 Health Systems Strengthening: Health Systems 2020. Chapter 11: Health Information System Module 11.1. Web material, accessed on 25 Oct. 14, at: www.healthsystems2020.org/files/581_file_11_Chapter_11.pdf

51 International Monetary Fund. Heavily Indebted Poor Countries and Poverty Reduction Strategy Paper (HIPC/PRSP). Information available at https://www.imf.org/external/np/sec/nb/2000/nb0079.htm visited on Nov 1, 2014.

52 The U.S. President's Emergency Plan for AIDS Relief (PEPFAR). Information available at http://www.pepfar.gov/ accessed on Nov 1, 2014

53 The 3 by 5 Initiative. Information available at http://www.who.int/3by5/en/ visited on Nov 1, 2014.

54 The Global Fund to fight against AIDS, Tuberculosis, and Malaria. Information available at: http://www.theglobalfund.org/en/ accessed on Nov 1, 2014

55 The Rollback Malaria (RBM). Information available at: http://www.rollbackmalaria.org/rbmmandate.html accessed on Nov 1, 2014

56 Stop TB Partnership. Information available at: http://www.stoptb.org/ accessed on Nov 1, 2014

57 The Safe Motherhood Initiative. Information available at: http://www.safemotherhood.org/ accessed on Nov 1, 2014

58 Meessen B, Soucat A, Sekabaraga C. Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform? Bulletin of the World Health Organization 2011;89:153-156. doi: 10.2471/BLT.10.077339

59 Ireland M, Paul E, Dujardin B. Can performance-based financing be used to reform health systems in developing countries? Bulletin of the World Health Organization. Bull World Health Organ. 2011, vol.89 (9) http://dx.doi.org/10.1590/S0042-96862011000900017

60 The World Bank. A brief note on Monitoring and Evaluation from the Funder's perspective. Information available at:

http://siteresources.worldbank.org/INTBELARUS/Resources/M&E.pdf accessed on Nov 1, 2014.

61 Chan M, Kazatchkine M, Lob-Levyt J, Obaid T, Schweizer J, et al. (2010) Meeting the Demand for Results and Accountability: A Call for Action on Health Data from Eight Global Health Agencies. PLoS Med 7(1): e1000223.

62 International Health Partnership and related initiatives (IHP+). Monitoring and Evaluation Lessons. Available at: http://www.internationalhealthpartnership.net/en/results-evidence/me-lessons/ . Accessed on Nov 1 2014.

63 Monitoring, evaluation and review of national health strategies: a country-led platform for information and accountability. World Health Organization. 2011. Available at: http://www.internationalhealthpartnership.net/fileadmin/uploads/ihp/Documents/Tools/M_E_Framework/M%26E.framework.2011.pdf. Accessed Nov 1 2014.

64 Habicht JP, Victora CG and Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. Int Jour Epi. 1999; 28:10 – 18

65 Kirkwood BR, Cousens SN, Victora CG, Zoysa I. issues in the design and interpretation of studies to evaluate the impact of communit-based interventions. Trop Med Int Health 1997; 2:1022 – 29

66 Victora CG, Black RE, Boerma JT, Bryce J. Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations. 2011; 377 (9759): 85-95.

67 Stoltzufus RJ, Pillai G. Measuring performance: a strategy to improve programs. J Nutrition. 2002 Apr;132(4 Suppl):845S-8S.

68 Victora CG. Habicht JP, Bryce J. Evidence-Based Public Health: Moving Beyond Randomized Trials. Am J Public Health. 2004 March; 94(3): 400–405.

69 Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. Framework for design and evaluation of complex interventions to improve health. BMJ 2000; 321: 694–6.

70 Habicht JP, Martorell R. Probability, plausibility and adequacy evaluations of the Oriente Study demonstrate that supplementation improved child growth. J. Nutrition. 2010. 140: 407 – 410

71 Rothwell PM. External validity of randomised controlled trials: To whom do the results of this trial apply? Lancet. 2005;365:82–93.

72 Rothwell PM (2006) Factors that can affect the external validity of randomised controlled trials. PLoS Clin Trials 1(1): e9. DOI: 10.1371/journal.pctr.0010009

73 Duignan Paul. Evaluation types: Formative/developmental, process and impact/outcome: A topic article in the Outcomes Theory Knowledge Base [Internet, accessed on 16th Oct 2014]. Version 1. Outcomes Theory Knowledge Base. 2009 Sep 15. Available from: http://outcomestheory.wordpress.com/article/evaluation-types-formative-2m7zd68aaz774-119/

74 Livingston J, Reback D. Process Evaluation. Aging and Disability Resource Center. Evaluation Peer Workgroup. Flint Springs Associates. Workshop presentation, July 26, 2007. Internet, accessed on 26th Oct 2014, at www.adrc-tae.acl.gov/tiki-download_file.php?fileId=26630

75 Dusenbury L, Brannigan R, Falco M, Hansen W: A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. Health Educ Res 2003,18:237-256

76 Pawson R and Tilley N. Realist Evaluation.2004. accessed on 16th Oct 2014, document available at: www.communitymatters.com.au/RE_chapter.pdf

77 Lhussier M, Bangash S, Dykes F, Zaman M, Lowe NM. Development and implementation of a nutrition intervention programme in North West Pakistan: a realist framework. Health Promotion International, 2011. Vol. 27 No. 4 doi:10.1093/heapro/dar067

78 Clark H, Taplin D. Theory of Change Basics: A Primer on Theory of Change. 2012. New York: Actknowledge. Book accessed on Internet on 26th Oct 2014, at: http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/ToCBasics.pdf

79 Anderson, A. (2005). The community builder's approach to theory of change: A practical guide to theory and development. New York: The Aspen Institute Roundtable on Community Change.

80 Vogel I: Review of the use of Theory of Change in International Development. London: UK Department of International Development; 2012.

81 James C: Theory of Change Review: A report comissioned by Comic Relief; 2011. Accessed on Nov 3, 2011, at: http://www.theoryofchange.org/pdf/James_ToC.pdf

82 Measuring the strength of implementation of community case management of childhood illness within the Catalytic Initiative to Save a Million Lives. Working Paper Version 27 August 2011, paper available for download at: http://www.jhsph.edu/departments/international-health/centers-and-institutes/institute-for-international-programs/_documents/rapid_scaleup/wp-implementation-strength.pdf

83 Corday, D. S.,&Pion, G. M. (1993). Psychological rehabilitation assessment: A broader perspective. In R. L. Flueckauf, L. B. Sechrest, G. R. Bond, & E. C. McDonel (Eds.), Improving assessment in rehabilitation and health (pp. 215–240). Thousand Oaks, CA: Sage.

84 Miller NP, Amouzou A, Tafesse M, Hazel E, Legesse H, Degefie T, Victora CG, Black RE, Bryce J. Integrated Community Case Management of Childhood Illness in Ethiopia: Implementation Strength and Quality of Care. Am. J. Trop. Med. Hyg., 91(2), 2014, pp. 424–434 doi:10.4269/ajtmh.13-0751

85 Institute for International Programs, Johns Hopkins University, National Statistical Office Zomba, Malawi, Save the Children Lilongwe, Malawi and Washington DC, USA. CCM Implementation Strength Measurement: Validation of Mobile Phone Interviews with Health Surveillance Assistants in Malawi. 2013. Report accessed on 27th Oct 2014, at: http://www.jhsph.edu/departments/international-health/centers-and-institutes/institute-for-international-programs/projects/traction/Malawi_IS_CVS_report_final.pdf

86 Scott AG and Sechrest L. Strength of theory and theory of strength. Eval and Progr Planning. 1989; 12:4, 329 – 336

87 Yeaton WH & Sechrest L. Critical Dimensions in the Choice and Maintenance of Successful Treatments: Strength, Integrity, and Effectiveness. Journal of Consulting and Clinical Psychology. 1981, Vol. 49, No. 2, 156-167

88 Tanasescu M, Leitzmann MF, Rimm EB, Willett WC, Stampfer MJ, Hu FB. Exercise type and intensity in relation to coronary heart disease in men. JAMA. 2002 Oct 23-30;288(16):1994-2000.

89 Graham WJ, Cairns J, Bhattacharya S, et al. Disease control priorities in developing countries, 2nd ed. New York: Oxford University Press, 2006: 499–529.

90 Bhutta ZA, Das JK, Bahl R, Lawn JE, Salam RA, Paul VK, Sankar MJ, Blencowe H, Rizvi A, Chou VB, Walker N, for The Lancet Newborn Interventions Review Group, The Lancet Every Newborn Study Group. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? The Lancet, Volume 384, Issue 9940, Pages 347 - 370, 26 July 2014 doi:10.1016/S0140-6736(14)60792-3r

91 Jamison DT, Breman JG, Measham AR, et al. Disease control priorities in developing countries, 2nd Edn. New York: Oxford University Press, 2006.

92 Lyman GH. Impact of Chemotherapy Dose Intensity on Cancer Patient Outcomes J Natl Compr Canc Netw 2009;7:99-108

93 Waltz J, Addis ME, Koemer K, Jacobson NS. Testing the fidelity of a psychotherapy protocol: assessment of adheherence and competence. Journal of Consulting and Clinical Psyc. 1993. 61, 620 – 630

94 London School of Hygiene and Tropical Medicine, The IDEAS project. Measuring implementation strength. Literature review draft report 2012. Accessed on 1st Sep 2013, available at: http://ideas.lshtm.ac.uk/sites/ideas.lshtm.ac.uk/files/Report_implementation_strength_Final_0.pdf

95 McGrew JH, Bond GR, Dietzen L, Salyers M. Measuring the fidelity of implementation of a mental health program model. Journal of Consult Clin Psychol. 1994 Aug; 62(4):670-8.

96 Ross J, Stover J. The Family Planning Program Effort Index: 1999 Cycle. Int Family Plan Pers. 2001;Vol 27 (3)

97 Ross J and Stover J. Effort Indices for National Family Planning Programs, 1999 Cycle. The Futures Group International, 2000

98 Bulatao RA, Ross JS. Rating maternal and neonatal health services in low- and middle-income countries. Bulletin of the World Health Organization 2002; 80:721-727.

99 Sicsic Jr, Le Vaillant M, Franc C. Building a composite score of general practitioners' intrinsic motivation: a comparison of methods. Int J Qual Health Care. 2014 Apr;26(2):167-73. doi: 10.1093/intqhc/mzu007. Epub 2014 Feb 11.

100 Jacobs R, Smith P, Goddard M. Measuring performance: An examination of composite performance indicators. University of York, Centre for Health Economics. CHE Technical Paper Series 29, Web, 20th Nov 2014, available at: http://www.york.ac.uk/che/pdf/tp29.pdf

101 Srivastava N, Sandhu A. Infant and child feeding index. Indian J Pediatr. 2006 Sep;73(9):767-70.

102 Antony GM, Rao KV. A composite index to explain variations in poverty, health, nutritional status and standard of living: use of multivariate statistical methods. Public Health. 2007 Aug;121(8):578-87. Epub 2007 Apr 30.

103 Sicsic J, Le Vaillant M, Franc C. Building a composite score of general practitioners' intrinsic motivation: a comparison of methods. Int J Qual Health Care. 2014 Apr;26(2):167-73. doi: 10.1093/intqhc/mzu007. Epub 2014 Feb 11.

104 Foa R, Tanner JC. Methodology of the Indices of Social Development. Web content retrieved on 10.1.2015 from http://www.indsocdev.org/resources/Methodology%20of%20the%20Social%20Development%20Indices_%20jan11.pdf

105 Organisation for Economic Cooperation and Development. Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD and JRC-European Commission. 2008. [web content, accessed on 23rd Nov 2013, available at www.oecd.org/std/42495745.pdf]

106 Otieno GO, Hinako T, Motohiro A, Daisuke K, Keiko N. Measuring effectiveness of electronic medical records systems: towards building a composite index for benchmarking hospitals. Int J Med Inform. 2008 Oct;77(10):657-69. doi: 10.1016/j.ijmedinf.2008.01.002. Epub 2008 Mar 4.

107 Ross JA, Campbell OMR, Bulatao R. The Maternal and Neonatal Programme Effort Index (MNPI). Tropical Medicine and International Health. 2001. 6:10 pp. 787 -798

108 The Economist Intelligence Unit. Quality-of-life index. The world in 2005. [web content, accessed on 24th Nov 2013, available at: www.economist.com/media/pdf/QUALITY_OF_LIFE.pdf]

109 Villar J, Ba'aqeel H, Piaggio Ga, Lumbiganon P, Belizán JM, Farnot U, et al (for the WHO Antenatal Care Trial Research Group 2001). WHO antenatal care randomised trial for the evaluation of a new model of routine antenatal care. The Lancet. 2001. Vol 357 9268, 1551 – 1564.

110 Carroli G, Villar J, Piaggio G, Khan-Neelofur D, Gulmezoglu M, Mugford M, Lumbiganon P, Farnot U, Bersgjo P for the WHO Antenatal Care Trial Research Group. WHO systematic review of randomised controlled trials of routine antenatal care. The Lancet. 2001, Vol 357. Issue 9268, 1565 – 1570

111 UNDP/UNFPA/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction. Department of Reproductive Health and Research. Family and Community Health. WHO Antenatal Care Randomized Trial: Manual for the implementation of the new model. Geneva: WHO, 2002.

112 Armar-Klemesu, M. et al. 2006. "Acceptability and feasibility of introducing the WHO focused antenatal care package in Ghana," FRONTIERS Final Report. Washington, DC: Population Council.

113 United Republic of Tanzania. Ministry of Health and Social Welfare/Reproductive and Child Health Section/National Malaria Control Programme. Focused Antenatal Care. Malaria and Syphilis during Pregnancy. Orientation Package for Service Providers. Dar es Salaam: Ministry of Health and Social Welfare, 2004.

114 Kinzie B and Gomez P. Basic maternal and newborn care: A guide for skilled providers. 2004. Jhpiego/MNH program: Baltimore Maryland.

115 Gross K, Schellenberg JA, Kessy F, Pfeiffer C, Obrist B. Antenatal care in practice: an exploratory study in antenatal care clinics in the Kilombero Valley, south-eastern Tanzania. BMC Pregnancy and Childbirth 2011, 11:36 doi:10.1186/1471-2393-11-36

116 Dowswell T, Carroli G, Duley L, Gates S, Gülmezoglu AM, Khan-Neelofur D, Piaggio GGP. Alternative versus standard packages of antenatal care for low-risk pregnancy (Review). The Cochrane Library 2011, Issue 4.

117 Pembe AB, Carlstedt A, Urassa DP, Lindmark G, Nystrom L, Darj E. Quality of antenatal care in rural Tanzania: counselling on pregnancy danger signs. BMC Pregnancy and Childbirth 2010, 10: 35

118 Urassa DP, Carlstedt A, Nystrom L, Massawe SN, Lindmark S. Quality assessment of the antenatal program for anaemia in rural Tanzania. Int J Qual Health Care. 2002. 14 (6): 441-448.

119 von Both C, Flessa S, Makuwani A, Mpembeni R, Jahn A. How much time do health services spend on antenatal care? Implications for the introduction of the focused antenatal care model in Tanzania. BMC Pregnancy and Childbirth. 2006, 6: 22

120 Boller C, Wyss K, Mtasiwa D, Tanner M. Quality and comparison of antenatal care in public and private providers in the United Republic of Tanzania. Bulletin of the World Health Organization 2003; 81: 116-122

121 Rockers PC, Wilson ML, Mbaruku G, Kruk ME. Source of antenatal care influences facility delivery in rural Tanzania: a population-based study. Matern Child Health J. 2009. 13 (6) 879 - 885

122 Magoma M, Requejo R, Campbell OMR, Cousens S, Filippi V. High ANC coverage and low skilled attendance in a rural Tanzanian district: a case for implementing a birth plan intervention. BMC Pregnancy Childbirth. 2010; 10: 13

123 World Health Organization. Essential obstetric care. Web material accessed on 29th Apr 2014, available at: http://www.who.int/mediacentre/factsheets/fs245/en/

124 Mangham-Jefferies L, Pitt C, Cousens S, Mills A, Schellenberg J. Cost-effectiveness of strategies to improve the utilization and provision of maternal and newborn health care in low-income and lower-middle-income countries: a systematic review. BMC Pregnancy Childbirth. 2014 Jul 22;14(1):243. [Epub ahead of print]

125 World Health Organization and UNICEF, "Building a Future for Women and Children: The 2012 Report.

126 Adanu R. The challenge of meeting the Millennium Development Goal for maternal health. Int J Gynecol Obstet 2008;102:1–2.

127 Kongnyuy E, Hofman J, van den Broek N. Ensuring effective Essential Obstetric Care in resource poor settings. BJOG 2009;116 (Suppl. 1):41–47

128 WHO, UNFPA, UNICEF, AMDD. Monitoring emergency obstetric care: a handbook. France, WHO, 2009.

129 UNICEF, WHO, UNFPA. Guidelines for monitoring the availability and use of obstetric services. New York. UNICEF, 1997

130 Thaddeus S, Maine D. Too far to walk: maternal mortality in context. Soc Sci Med. 1994 Apr; 38(8):1091-110.

131 Donabedian A. The Quality of Care How Can It Be Assessed? JAMA 1988;260:1743-1748

132 Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. Implementation Science 2007, 2:40 doi:10.1186/1748-5908-2-40

133 World Health Organization. Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies. WHO, Geneva. 2010.

134 de Savigny D, Adam T (Eds). Systems thinking for health systems strengthening. Alliance for Health Policy and Systems Research, WHO, 2009.

135 Last J. 2001. Dictionary of epidemiology. 4 ed. Oxford: Oxford University Press.

136 Victora GC , Schellenberg JA, Huicho L, Amaral J, El Arifeen S, Pariyo G, Manzi F, Scherpbier RW, Bryce J, Habicht JP. Context matters: interpreting impact findings in child survival evaluations. Health Policy Plan. (December 2005) 20 (suppl 1): i18-i31. doi: 10.1093/heapol/czi050

137 World Health Organization. Health statistics and Health Information Systems. Methods, indicators and questionnaire for the Service Availability and Readiness Assessment survey available at http://www.who.int/healthinfo/systems/sara_introduction/en/index.html. Accessed in November 2012.

138 World Health Organisation. Accountability for Women's and Children's Health. Recommendations for health indicators. Internet access, Nov 4th 2011, at: http://www.who.int/woman_child_accountability/progress_information/recommendation2/en/

139 Becker L, Pickett J, R L: Measuring commitment to health: Global health indicators working group report. [www.cgdev.org/content/publications/detail/10016] Washington, DC: Center for Global Development; 2006; (Accessed 04 Nov 2014)

140 WHO Department of Reproductive Health and Research. "Technical Consultation on Reproductive Health Indicators 21-22 September 2005 Meeting Report." Geneva, 2005. World Health Organization.

141 Summerfelt WT. Program strength and fidelity in evaluation. Applied Devel. Scie. 2003. Vol 7, No 2, 55 – 61.

142 El Arifeen S, Blum L, Hoque D, Chowdhury E, Khan R, Black RE, et al., et al. Integrated Management of Childhood Illness (IMCI) in Bangladesh: early findings from a cluster-randomised study. Lancet 2004; 364: 1595-602 doi: 10.1016/S0140-6736(04)17312-1 pmid: 15519629.

143 The World Health Organisation. www.who.int. Visited on 26th Dec 2014

144 Jpiego – An affiliate of the Johns Hopkins University. www.jpiego.org. Visited on 26th Dec 2014

145 Lewis SM, Cullinane FM, Carlin JB, Halliday JL. Women's and health professionals' preferences for prenatal testing for Down syndrome in Australia. Aust N Z J Obstet Gynaecol. 2006 Jun;46(3):205-11.

146 Bekker-Grob EW, Essink-Bot ML, Meerding WJ, Koes BW, Steyerberg EW. Preferences of GPSs and patients for preventive osteoporosis drug treatment: a discrete-choice experiment. Pharmacoeconomics. 2009: 27 (3):211-9

147 Bishop AJ, Marteau TM, Armstrong D, Chitty LS, Longworth L, Buxton MJ, Berlin C. Women and health care professionals' preferences for Down's Syndrome screening tests: a conjoint analysis study. BJOG. 2004 Aug;111(8):775-9.

148 The District Health Information Software version 2, available at www.dhis2.org. Last visited on 29th Jan 2014

149 World Health Organization. "Achieving the health-related MDGs. It takes a workforce". Web materials accessed on 29th Jan 2014 at: www.who.int/hrh/workforce_mdgs/en/

151 Abate M, Diegert K, Allen H. A hierarchical approach to improving data quality. Data Qual J. 1998;33(4):365–9.

152 Personal communication with MOHSW staff who emailed the author the Introduction Chapter and the Official Statistics "ANNUAL HEALTH STATISTICAL TABLES AND FIGURES, MAINLAND TANZANIA, HMIS 2012"

153 Tanzanian Ministry of Planning, Economy and Empowerment. National Bureau of Statistics. Tanzania population projections 2003 to 2025. Popular version, Dar es Salaam, February 2008. Soft copy was availed from the National Bureau of Statistics.

154 United Nations Development Fund. Human Development Index. Human Development Index Reports. Webmaterial accessed on 31st Jan 2014 at: http://hdr.undp.org/en/statistics/hdi

155 Acuna E. and Rodriguez C., The treatment of missing values and its effect in the classifier accuracy. In D. Banks, L. House, F.R. McMorris,P. Arabie, W. Gaul (Eds).Classification, Clustering and Data Mining Applications. Springer-Verlag Berlin-Heidelberg, 639-648. 2004. [comment has been widely cited by papers on multiple imputations/missing data]

156 United Republic of Tanzania. Bank of Tanzania: Financial Markets - Indicative Foreign Exchange Market Rates. Date on 31st Dec 2010 as indicative middle date for Jan 2010 – Dec 2011, spot selling rate of 1 GBP=2,267 TZS and 1 USD=1,468. Accessed in April 2013, available at: http://bot.go.tz/FinancialMarkets/FinancialMarkets.asp#ForeignExchangeMarkets

157 The World Health Organization. The human resources for health crisis. 2014. [Web material] Accessed 25.04.2014 at: http://www.who.int/workforcealliance/about/hrh_crisis/en/

158 Make Motherhood Safe for Tanzanian Women. Comprehensive Community Based Rehabilitation in Tanzania (CCBRT), Dar es Salaam. Web, 18th Nov 2014, at: http://www.globalgiving.org/projects/one-bed-per-mother/

159 Pitchforth E, Lilford RJ, Kebede Y, Asres G, Stanford C, Frost J. Assessing and understanding quality of care in a labour ward: A pilot study combining clinical and social science perspectives in Gondar, Ethiopia. Social Science & Medicine 71 (2010) 1739e1748

160 Mayi-Tsonga S, Oksana L, Ndombi I, Diallo T, de Sousa MH, Faúndes A. Delay in the provision of adequate care to women who died from abortion-related complications in the principal maternity hospital of Gabon. Reproductive Health Matters 2009;17(34):65–70

161 National Bureau of Statistics (NBS) [Tanzania] and ICF Macro. 2011. Tanzania Demographic and Health Survey 2010. Dar es Salaam, Tanzania: NBS and ICF Macro. Report accessed in June 2013 and can be downloaded from: http://www.measuredhs.com/publications/publication-fr243-dhs-final-reports.cfm

162 World Health Organisation. Indicators for Monitoring Millennium Development Goals: Definitions, Rationale, Concepts and Sources for Antenatal Care Coverage. Web, visited on 26th Dec 2014, at: http://mdgs.un.org/unsd/mi/wiki/5-5-Antenatal-care-coverage-at-least-one-visit-and-at-least-four-visits.ashx

163 The United Republic of Tanzania. Ministry of Health and Social Welfare. Integrated Logistics System (ILS). Procedures Manual Roll-Out Version. September, 2008. Accessed on Nov 18th, 2013, at: http://nexus.som.yale.edu/ph-tanzania/sites/nexus.som.yale.edu.ph-tanzania/files/imce_imagepool/ILS_Rollout_Rolesand%20more_2008.pdf

164 Chimnani J, Kamunyori J, Bancroft E, Nazarewicz N, Kisoka N, Mahamba J. 2010. Tanzania: Review of the Health Facility Report and Request Forms at MSD Zonal Stores. Arlington, Va.: USAID | DELIVER PROJECT, Task Orders 1 and 3.

165 McIntyre, Diane. "Beyond Fragmentation and towards Universal Coverage: Insights from Ghana, South Africa and the United Republic of Tanzania." Bulletin of the World Health Organization. WHO, Nov. 2008. Web. 04 Jan. 2013.

166 Mills A, Ataguba JE, Akazili J, Borghi J, Garshong B, Makawia S, Mtei G, Harris B, Macha J, Meheus F, McIntyre D. Equity in financing and use of health care in Ghana, South Africa, and Tanzania: implications for paths to universal coverage. Lancet. 2012 Jul 14;380(9837):126-33. doi: 10.1016/S0140-6736(12)60357-2. Epub 2012 May 15.

167 Macha J, Harris B, Garshong B, Ataguba JE, Akazili J, Kuwawenaruwa A, Borghi J. Factors influencing the burden of health care financing and the

distribution of health care benefits in Ghana, Tanzania and South Africa. Health Policy Plan. 2012 Mar;27 Suppl 1:i46-54.

168 Kruk, Margaret E., Godfrey Mbaruku, Peter C. Rockers, and Sandro Galea. "User Fee Exemptions Are Not Enough: Out-of-pocket Payments for 'free' Delivery Services in Rural Tanzania." Tropical Medicine & International Health 13.12 (2008): 1442-451. Print.

169 World Health Organisation. Global Health Expenditure Database. http://apps.who.int/nha/database/DataExplorerRegime.aspx . Visited on 26th Dec 2014

170 The Abuja Declaration 2001, www.un.org/ga/aids/pdf/abuja_declaration.pdf. Retrieved on 26th Dec 2014

171 The World Bank, Africa Region. Human Development Department. Making Health Financing Work for Poor People in Tanzania. Health Financing Policy Note. Sep 2011. Web, 18th Nov 2014, at http://p4h-network.net/wp-content/uploads/2013/10/WB_TanzaniaHealthFinancingPolicyNoteFinal.pdf

172 World Health Organization: Training manual on management of human resources for health. Section I, part A. Geneva 1993. Visited on 23rd Nov 2013

173 Zhou P, Ang BW, Zhou DQ. Weighting and aggregation in composite index construction: A multiplicative optimization approach. Social Indicators and Research. 2010. 96. pp 169 – 181

174 McGraw Hill Financial. S&P Dow Jones Indices. Dow Jones Industrial Average History. [Web content, visited on 24th Nov 2013, at http://eu.spindices.com/about-sp-indices/our-history/]

175 Boskin MJ. Consumer Price Indexes. The Concise Encyclopedia of Economics. David R. Henderson, ed. Liberty Fund, Inc. 2008. Library of Economics and Liberty [Online] available from http://www.econlib.org/library/Enc/ConsumerPriceIndexes.html; accessed 24 November 2013; Internet.

176 Anand S, Hanson K. Disability-adjusted life years: a critical review. Journal of Health Economics. 1999,Vol 16, Issue 6, pp 685–702, available at: http://dx.doi.org/10.1016/S0167-6296(97)00005-2

177 Evans DB, Tandon A, Murray CJL, Lauer J: The comparative efficiency of national health systems in producing health: an analysis of 191 countries. In Global Programme on Evidence for Health Policy Discussion Paper No. 29. Geneva, Switzerland, World Health Organization; 2000.

178 Tsui A. Monitoring National Progress with Composite Indices. In: Measure Evaluation. Monitoring population and health program efforts with composite indices. MEASURE Evaluation Bulletin. 2001, Number 3. [web content, accessed in Nov 2013], available at: http://pdf.usaid.gov/pdf_docs/pnacn324.pdf

179 Ben-Hur and I. Guyon. Detecting stable clusters using principal component analysis. In M.J. Brownstein and A. Kohodursky, editors, Methods In Molecular Biology, pages 159–182. Humana Press, 2003.

180 Jolliffe, I. (1972) Discarding variables in principal component analysis I: Artificial data. Applied Statistics, 21, 160–173.

181 Jollife, IT. Principal Component Analysis. New York, 2002, Second Edition. Springer. Book available in pdf for download at: http://f3.tiera.ru/2/M_Mathematics/MV_Probability/MVsa_Statistics%20and%20applications/Jolliffe%20I.%20Principal%20Component%20Analysis%20%282ed.,%20Springer,%202002%29%28518s%29_MVsa_.pdf

182 Vyas S, Kumarnayake L. Constructing socio-economic status indices: how to use principal components analysis. Health Policy Plan. 2006;21:459 – 68

183 Freudenberg M. Composite indicators of country performance: a critical assessment, OECD, 2003. Paris.

184 Thabane L, Mbuagbaw L, Zhang S, Samaan Z, Marcucci M, Ye C, Thabane M, Giangregorio L, Dennis B, Kosa D, Debono VB, Dillenburg R, Fruci V, Bawor M, Lee J, Wells G, Goldsmith GH. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. BMC Medical Research Methodology 2013, 13:92 doi:10.1186/1471-2288-13-92

185 Gouws E, Bryce J, Pariyo G, Armstrong Schellenberg J, Amaral J, Habicht JP. Measuring the quality of child health care at first-level facilities. Soc Sci Med. 2005 Aug;61(3):613-25. Epub 2005 Feb 23.

186 Colbourn T, Masache G, Skordis-Worrall. Development, reliability and validity of the Chichewa WHOQOL-BREF in adults in Lilongwe, Malawi. BMC Research Notes 2012, 5:346

187 Strainer DL and Norman GR. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press, 2008. Fourth Edition

188 Ponterotto IG, Ruckdeschel D.An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. Perceptual and Motor Skills, 2007. 105, pp 997 – 1014.

189 Ross JA, Begala JE. Measures of Strength for Maternal Health Programs in 55 Low- and middle-income countries: The MNPI Study. Maternal and Child Health Journal, Vol. 9, No. 1, March 2005 (C _ 2005) DOI: 10.1007/s10995-005-2548-z

190 Smith, P. (2002) Developing composite indicators for assessing health system efficiency, in Smith, P.C. (ed.) Measuring up: Improving the performance of health systems in OECD countries, OECD: Paris.

191 Roseboom WW. The reliability of a linear composite of non-equivalent subtests. Applied Psychological Measurement. 1989. 13:3, pp. 277 – 283.

192 Hagerty MR, Land KC. Constructing summary indices of quality of life: A model for the effect of heterogenous importance weights. Sociological Methods and Research. 2007. 35, pp 455 – 496.

193 World Health Organization. Country profiles of environmental burden of disease, World Health Organization. 2009. Access: www.who.int/quantifying_ehimpacts/countryprofiles. Visited on 24th Nov 2013

194 The Economist Intelligence Unit. Best cities to be born in 2013. [web content, accessed on 24th Nov 2013, at: http://www.eiumedia.com/index.php/daily-comment/item/851-best-country-to-be-born-in-2013-announced]

195 Crane PK, Carle A, Gibbons LE, Insel P, Mackin RS, Gross A, Jones RN, Mukherjee S, Curtis SM, Harvey D, Weiner M, Mungas D; Alzheimer's Disease Neuroimaging Initiative. Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). Brain Imaging Behav. 2012 Dec;6(4):502-16. doi: 10.1007/s11682-012-9186-z.

196 World Health Organization, Department of Reproductive Health and Research. 2003. Antenatal care in developing countries: promises, achievements and missed opportunities: an analysis of trends, levels and differentials, 1990-2001

197 Lilford RJ, Chard T. Problems and pitfalls of risk assessment in antenatal care. Br J Obstet Gynaecol. 1983 Jun;90 (6):507-10.

198 Yuster EA. Rethinking the role of the risk approach and antenatal care in maternal mortality reduction. Int J of Gyn & Obst. 1995. Vol 50, Supp 2, Pages S59–S61

199 Villar J, Ba'aqeel H, Piaggio G, Lumbiganon P, Miguel Belizán J, Farnot U, Al-Mazrou Y, Carroli G, Pinol A, Donner A, Langer A, Nigenda G, Mugford M, Fox-Rushby J, Hutton G, Bergsjø P, Bakketeig L, Berendes H, Garcia J; WHO Antenatal Care Trial Research Group. WHO antenatal care randomised trial for the evaluation of a new model of routine antenatal care. Lancet. 2001 May 19;357(9268):1551-64. Erratum in: Lancet 2001 Nov 3;358(9292):1556.

200 Munjanja SP, Lindmark G, Nyström L. Randomised controlled trial of a reduced-visits programme of antenatal care in Harare, Zimbabwe. Lancet. 1996 Aug 10; 348(9024):364-9.

201 Basic maternal and newborn care: a guide for skilled providers (reference manual learning resource package), 2004. Jhpiego/MNH program; Jhpiego/Maternal and Neonatal Health Program, 2001. Focused antenatal care: Planning and providing care during pregnancy. Jhpiego, Baltimore, MD

202 Malecela, M., W. Kisinza, W. Kisoka, M. Kamugisha, P. Mutalemwa, W. Makunde, V. Barongo, J. Urassa, G. Kimaro, D. Urassa, K. Senkoro, G. Mfinanga, M. Mayige, E. Charles, S. Kibona, J. Massaga, G. Nkya, U. Mwingira, M. Munga, Lugimbana, C. Kihinga, C. Sanga, E. Mkondya-Senkoro, R. Kigadye, L. Mfalila, T. John and A. Kitua (2006). Situation analysis of emergency obstetric care for safe motherhood in public health facilities in mainland Tanzania. Dar es Salaam: National Institute of Medical Research and Ministry of Health.

203 Olsen ØE, Ndeki S, Norheim OF. Availability, distribution and use of emergency obstetric care in northern Tanzania. Health Policy Plan. 2005 May;20(3):167-75.

204 Jhpiego – an Affiliate of Johns Hopkins University. Tanzania's country profile. Web, 28th Nov 2014, at: http://www.jhpiego.org/content/tanzania

205 Shija AE, Msovela J, Mboera LEG. Maternal health in fifty years of Tanzania independence: Challenges and opportunities of reducing maternal mortality. Tanzania Journal of Health Research. 2011. Vol. 13 (Suppl 1). DOI:http://dx.doi.org/10.4314/thrb.v13i5.5.

206 United Republic of Tanzania Ministry of Health and Social Welfare. The National Road Map Strategic Plan To Accelerate Reduction of Maternal, Newborn and Child Deaths in Tanzania: 2008 – 2015. April 2008. Retrieved on 12th Jan 2015 from: http://www.who.int/pmnch/countries/tanzaniamapstrategic.pdf

207 Lincetto O, Mothebesoane-Anoh S, Gomez P, Munjanja SP: Antenatal care. In Opportunities for Africa's newborns: Practical data, policy and programmatic support for newborn care in Africa. Edited by Joy L, Kate K. Cape Town: PMNCH; 2006. Web, 28th Nov 2014, at: http://www.who.int/pmnch/media/publications/africanewborns/en/

209 Hellerstein, Joseph M. "Quantitative data cleaning for large databases." United Nations Economic Commission for Europe (UNECE) (2008). Web, 29th Nov. 2014, at: http://db.cs.berkeley.edu/jmh/papers/cleaning-unece.pdf

210 Dasu T, Johnson T. Exploratory Data Mining and Data Cleaning. Wiley, 2003

211 AbouZahr C, Wardlaw T. Antenatal Care in Low- and middle-income countries: Promises, Achievements and Missed Opportunities: An Analysis of Trends, Levels and Differentials, 1990-2001". Geneva, 2003

212 Tran TK, Nguyen CTK, Nguyen HD, Eriksson B, Bondjers G, Gottvall K, Ascher H, Petzold M. Urban - rural disparities in antenatal care utilization: a study of two cohorts of pregnant women in Vietnam. BMC Health Services Research 2011, 11:120

213 Pearson L, Shoo R. Availability and use of emergency obstetric care services: Kenya, Rwanda, Southern Sudan and Uganda. Int J Gynaecol Obstet. 2005; 88:208-215

214 Fauveau V; UNFPA – AMDD Working Group on Indicators. Program Note: Using UN process indicators to assess needs in emergency obstetric services: Gabon, Guinea-Bissau, and The Gambia. Int J Gynaecol Obstet. 2007; 96:233-240

215 The World Health Organisation. Maternal, Newborn, Child and Adolescent health. Indicators for Millennium Development Goals. Web, 29th Dec 2014, http://www.who.int/maternal_child_adolescent/topics/maternal/mdg/en/

216 National Bureau of Statistics (NBS) [Tanzania] and Macro International Inc.. 2007. Tanzania Service Provision Assessment Survey 2006. Dar es Salaam, Tanzania: National Bureau of Statistics and Macro International Inc. Full report available at: http://pdf.usaid.gov/pdf_docs/PNADL111.pdf

217 Tanzanian Ministry of Health and Social Welfare. Reproductive, Maternal, Newborn, and Child Health (RMNCH) Scorecard, web, visited on 31st July 2014 available at: http://www.mamaye.org/sites/default/files/evidence/Tanzania%20RMNCH%20scorecard.pdf

218 Say L and Raine R. A systematic review of inequalities in the use of maternal health care in low- and middle-income countries: examining the scale of the problem and the importance of context. Bull World Health Organ. 2007 October; 85(10): 812–819.

219 Mangham LJ, Hanson K. Scaling up in international health: what are the key issues? Health Policy and Planning 2010; 25:85–96

220 Arah OA, Klazinga NS, Delnoij DMJ, Ten Asbroek AHA, Custers T: Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement. International Journal for Quality in Health Care 2003,15:377-398

221 The United Nations Children's Fund. UNICEF data: Monitoring the situation of children and women – Definitions of indicators. Accessed on 9th Jan 2015, at: http://data.unicef.org/education/literacy

222 The World Bank. Country data on literacy rate, adult female. Accessed on 9th Jan 2015 at: http://data.worldbank.org/indicator/SE.ADT.LITR.FE.ZS

223 Deustche Well Swahili Service. Pictures of Dar es Salaam City floods. Taken on 22nd December 2011. Available at: https://www.facebook.com/dw.kiswahili/photos_stream [1st Aug 2013]

224 Tanzania's Daily Newspaper, Online Edition. "Will floods affect the city of Dar es Salaam again?" Article posted at http://www.dailynews.co.tz/index.php/features/popular-features/9184-will-floods-affect-the-city-of-dar-es-salaam-again [1st Aug 2013]

225 BBC News - Tanzania floods: Heavy rains inundate Dar es Salaam. Article posted at: http://www.bbc.co.uk/news/world-africa-16299734 [1st Aug 2013]

226 Alvarez JL, Gil R, Hernández V, Gil A. Factors associated with maternal mortality in Sub-Saharan Africa: an ecological study. BMC Public Health. 2009 Dec 14; 9:462

227 Koch E, Thorp J, Bravo M, Gatica S, Romero CX, Aguilera H, Ahlers I. Women's education level, maternal health facilities, abortion legislation and maternal deaths: a natural experiment in Chile from 1957 to 2007. PLoS One. 2012;7(5):e36613. doi: 10.1371/journal.pone.0036613. Epub 2012 May 4.

228 Chowdhury ME, Ahmed A, Kalim N, Koblinsky M. Causes of maternal mortality decline in Matlab, Bangladesh. Journal of Health Pop Nutr. 2009 Apr;27(2):108-23.

229 McTavish S, Moore S, Harper S, Lynch J. National female literacy, individual socio-economic status, and maternal health care use in sub-Saharan Africa. Soc Sci Med. 2010 Dec;71(11):1958-63.

230 Ollis S. Introduction to mHealth in Tanzania Improving access to high quality health care. D-tree International. Web, accessed on 5th Aug 2013, available at: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0CCwQFjAC&url=https%3A%2F%2Fwww.k4health.org%2Fsites%2Fdefault%2Ffiles%2FmHealth%2520in%2520TZ%2520-%2520Steve%2520Ollis.pptx&ei=uAaFVI6uFsi9Ub7GgQg&usg=AFQjCNGq6jrbMEtF_x-6wQri4-8AjeCPAQ&sig2=nRE0RCK37q50eNR1vNLyQQ

231 Barrington J, Wereko-Brobby O, Ward P, Mwafongo W, Kungulwe S. SMS for Life: a pilot project to improve anti-malarial drug supply management in rural Tanzania using standard technology. Malaria Journal 2010, 9:298

232 UNICEF Eastern and Southern Africa - Media Centre - Tanzania, 23 July 2013: The Government of Tanzania launches a new national birth registration system set to massively accelerate the number of children under 5 with birth certificates: [Web content, accessed on 5th Aug 2013], accessible at: http://www.unicef.org/esaro/5440_13078.html

233 Lund S, Nielsen BB, Hemed M, Boas ID, Said A, Said K, Makungu MH and Rasch V. Mobile phones improve antenatal care attendance in Zanzibar: a cluster randomized controlled trial. BMC Pregnancy and Childbirth 2014, 14:29 doi:10.1186/1471-2393-14-29

234 Africa Healthcare IT News. Tanzania is using Text messaging (SMS) service to boost maternity care. Web, 8th Dec 2014, at http://africahealthitnews.com/blogs/2013/04/tanzania-is-using-text-messaging-sms-service-to-boost-maternity-care/

235 Free C, Phillips G, Watson L, Galli L, Felix L, et al. (2013) The Effectiveness of Mobile-Health Technologies to Improve Health Care Service Delivery Processes: A Systematic Review and Meta-Analysis. PLoS Med 10(1): e1001363. doi:10.1371/journal.pmed.1001363

236 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 Med. 2011 Jan 25;8(1):e1000394. doi: 10.1371/journal.pmed.1000394.

237 United Republic of Tanzania. Ministry of Health and Social Welfare. Second Health Sector Strategic Plan (HSSP), July2003-June 2008, April 2003, Dar es Salaam.

238 Mpembeni RNM, Killewo JZ, Leshabari MT, Massawe SN, Jahn A, Mushi D, Mwakipa H. Use pattern of maternal health services and determinants of skilled care during delivery in Southern Tanzania: implications for achievement of MDG-5 targets. BMC Pregnancy and Childbirth 2007, 7:29

239 Watson JT, Gayer M, Connolly MA. Epidemics after Natural Disasters. Emerging Infectious Diseases, Jan 2007; 13(1): 1–5. doi: 10.3201/eid1301.060779

240 Gray CL, Mueller V. Natural disasters and population mobility in Bangladesh. Proc Natl Acad Sci USA. 2012 April 17; 109(16): 6000–6005

241 Marx M, Phalkey R, Guha-Sapir D. Integrated health, social, and economic impacts of extreme events: evidence, methods, and tools. Glob Health Action. 2012; 5: 10.3402/gha.v5i0.19837

242 Nour NN. Maternal Health Considerations During Disaster Relief. Rev Obstet Gynecol. 2011; 4(1): 22–27.

243 Costello A, Abbas M, Allen A, Ball S, Bell S, Bellamy R, et al. Managing the health effects of climate change: Lancet and University College London Institute for Global Health Commission. The Lancet. 2009;373:1693–733

244 Hill, Austin Bradford. The Environment and Disease: Association or Causation?. Proceedings of the Royal Society of Medicine. 1965:58 (5): 295–300. PMID 14283879

245 Block SR, Watkins SM, Salemi JL, Rutkowski R, Tanner JP, Correia JA, Kirby RS. Maternal pre-pregnancy body mass index and risk of selected birth defects: evidence of a dose-response relationship. Paediatr Perinat Epidemiol. 2013 Nov;27(6):521-31. doi: 10.1111/ppe.12084. Epub 2013 Oct 10.

246 Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health. 2005: 95 (Suppl 1): S144–50. PMID 16030331.

247 Karim AM, Admassu K, Schellenberg J, Alemu H, Getachew N, Ameha A, Tadesse L, Betemariam W. Effect of Ethiopia's health extension program on maternal and newborn health care practices in 101 rural districts: a dose-response study. PLoS One. 2013 Jun 4;8(6):e65160. doi: 10.1371/journal.pone.0065160. Print 2013.

248 Roberfroid D, Huybregts L, Habicht JP, Lanou H, Henry MC, Meda N, d'Alressandro U, Kolsteren P; MISAME Study Group. Randomized controlled

trial of 2 prenatal iron supplements: is there a dose-response relation with maternal hemoglobin? Am J Clin Nutr. 2011 May;93(5):1012-8. doi: 10.3945/ajcn.110.006239. Epub 2011 Mar 2.

249 Freeman K, Bonuck KA, Trombley M. Breastfeeding and infant illness in low-income, minority women: a prospective cohort study of the dose-response relationship. J Hum Lact. 2008 Feb;24(1):14-22; quiz 23-6. doi: 10.1177/0890334407310676.

250 Khatun S, Rahman M. Quality of antenatal care and its dose-response relationship with birth weight in a maternal and child health training institute in Bangladesh. J Biosoc Sci. 2008 May;40(3):321-37. Epub 2007 Nov 16.

251 Alba S, Nathan R, Schulze A, Mshinda H, Lengeler C. Child mortality patterns in rural Tanzania: an observational study on the impact of malaria control interventions. International Journal of Epidemiology.2014;43:204–215 doi:10.1093/ije/dyt231

252 Ross J, Smith E. 2010. The Family Planning Effort Index: 1999, 2004, and 2009. Washington, DC: Futures Group, Health Policy In itiative, Task Order 1. [Web, 15.12.2015], available at: http://www.healthpolicyinitiative.com/Publications/Documents/1110_1_FP_Effort_Index_1999_2004_2009__FINAL_05_08_10_acc.pdf

253 USAID, UNAIDS, WHO, and The Policy Project. The level of effort in the national response to HIV/AIDS: The AIDS program effort index (API) 2003 round. December 2003. Web, visited on 15th Dec 2014, available at: https://www.globalhivmeinfo.org/Gamet/pdf/619_The%20AIDS%20Program%20Effort%20Index,%202003.pdf

254 Gold RB, Singh S, Frost J. The Medicaid eligibility expansions for pregnant women: evaluating the strength of state implementation efforts. Fam Plann Perspect. 1993 Sep-Oct;25(5):196-207.

255 Gaber J, Gaber SL. Using face validity to recognize empirical community observations. Evaluation and Program Planning 33 (2010) 138–146

256 Nevo, B. (1985). Face validity revisited. Journal of Educational Measurement, 22(4), 287–293 Winter.

257 Secolsky, C. (1987). On the direct measurement of face validity: A comment on Nevo. Journal of Education Measurement, 24(1), 82–83 Spring.

258 Downing SM. Face validity of assessments: faith-based interpretations or evidence-based science? Med Educ 2006; 40:7-8

259 American Educational Research Association, American Pshychological Association, National Council on Measurement in Education. Standards for Education and Psychological Testing. Washington, DC: American Education Research Association, 1999

260 Dowming SM. Validity : on the meaningful interpretation of assessment data. Med Educ 2003;37:1-8

261 Anastasi, A. (1976). Psychological testing (fourth ed.). New York, NY: MacMillan Publishing.

262 Nunnally, J. (1978). Psychometric theory (second ed.). New York, NY: McGraw-Hill

263 Polit DF, Beck CT. The content validity index: Are you sure you know what's being reported? Critique and recommendations. Research in Nursing & Health, 2006. 29(5), 489–497.

264 Likert, R. "A Technique for the Measurement of Attitudes". Archives of Psychology, 1932, 140: 1–55.

265 Alwin DF, Krosnick JA. The reliability of survey attitude measurement: The influence of question and respondent attributes. Sociol. Methods Res. 1991, 20:139-81.

266 Polit DF, Beck C, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing & Health, 2007. 30(4), 459–467.

267 Bakeman, R.; Gottman, J.M. (1997). Observing interaction: An introduction to sequential analysis (2nd ed.). Cambridge, UK: Cambridge University Press. ISBN 0-521-27593-8.

268 Stata version 13. How to calculate Kappa statistics. Example 9 (number of raters >2 and number of ratings>2). Page 11. Retrieved on 13th Jan 2015 at: http://www.stata.com/manuals13/rkappa.pdf

269 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977, 33: 159–174.

270 Claeys C, Nève J, Tulkens PM, Spinewine A. Content Validity and Inter-Rater Reliability of an Instrument to Characterize Unintentional Medication Discrepancies. Drugs & Aging, 01 July 2012, vol./is. 29/7(577-591), 1170229X

271 Gibb H, Chapman-Novakofski K. Establishing content validity for the nutrition literacy assessment. Prev Chron Disorder. 2013. 10: 120267.

272 Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. Psycho Assess 1995;7(3):238 – 47

273 DeVellis RF. Scale development: theory and applications. 3rd edition. Los Angeles (CA): Sage Publications, 2012.

274 Neundorf, K. (2002). The content analysis guidebook. Thousand Oaks, CA: Sage Publications.

275 Underhill, N. (1987). Testing spoken language. Cambridge, UK: Cambridge University Press.

276 Gwet K. Handbook of Inter-Rater Reliability. 2014. Advanced Analytics Press, Maryland, USA. ISBN: 9 780970 806284.

277 Gwet K. Handbook of Inter-Rater Reliability. Gaithersburg, MD: STATAXIS Publishing Company; 2001. [Chapter 6 on Benchmarking Inter-Rater Reliability Coefficients, page 128]. Retrieved on 14th Jan 2015 at: http://www.agreestat.com/book3/bookexcerpts/chapter6.pdf

278 Huang KTL, Owino C, Vreeman RC, Hagembe M, Njuguna F, Strother RM, Gramelspacher GP. Assessment of the face validity of two pain scales in Kenya: a validation study using cognitive interviewing. BMC Palliative Care 2012, 11:5

279 Kelly DC, Margules AC, Kundavaram CR, Narins H, Gomella LG, Trabulsi EJ, Lallas CD. Face, content, and construct validation of the da Vinci Skills Simulator. Urology. 2012, 79 (5), 1068–1072

280 Varni JW, Curtis BH, Abetz LN, Lasch KE, Piault EC, Zeytoonjian AA. Content validity of the PedsQLTM 3.2 Diabetes Module in newly diagnosed patients with Type 1 diabetes mellitus ages 8–45. Qual Life Res, 2012. DOI 10.1007/s11136-012-0339-8

281 National Statistical Office, Malawi, 2011. Institute for International Programs—Johns Hopkins University, USA Measuring CCM Implementation Strength Report of Findings from a Pilot of Study Catalytic Initiative Rapid Scale-Up Program, Malawi. Lilongwe, Malawi: Government of Malawi.

282 Anderson, A. (2005). The community builder's approach to theory of change: A practical guide to theory and development. New York: The Aspen Institute Roundtable on Community Change.

283 Pawson R. Nothing as practical as a good theory. Evaluation. 2003, 9, 471–490.

284 Coryn CLS, Noakes LA, Westine CD, Schrote DC: A systematic review of theory-driven evaluation practice from 1990 to 2009. Am J Eval 2011, 32:199–226.

285 Weiss CH: Nothing As Practical As Good Theory: Exploring Theory-Based Evaluation For Comprehensive Community Initiatives For Children And Families. In New Approaches to Evaluating Community Initiatives Volume 1 Concepts, Methods and Contexts. Edited by Connell JP, Kubisch AC, Schorr LB, Weiss CH. Washington DC: The Aspen Institute; 1995:65–92.

286 De Silva MJ, Breuer E, Lee L, Asher L, Chowdhary N, Lund C, Patel V. Theory of Change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. Trials 2014, 15:267

287 Braveman PA. Monitoring Equity in Health and Healthcare: A Conceptual Framework. Journal of Hhealth Pop Nutri. 2003 Sep;21(3):181-192

288 Dingle A, Powell-Jackson T, Goodman C. A decade of improvements in equity of access to reproductive and maternal health services in Cambodia, 2000-2010. Int J Equity Health. 2013 Jul 9;12:51. doi: 10.1186/1475-9276-12-51.

289 Wilunda C, Putoto G, Manenti F, Castiglioni M, Azzimonti G, Erdessa W, Atzori A, Merialdi M, Betrán AP, Vogel J, Criel B. Measuring equity in utilization of emergency obstetric care at Wolisso Hospital in Oromiya, Ethiopia: a cross sectional study. Int J Equity Health. 2013 Apr 22;12:27. doi: 10.1186/1475-9276-12-27.

290 Victora CG, Barros AJ, Axelson H, Bhutta ZA, Chopra M, França GV, Kerber K, Kirkwood BR, Newby H, Ronsmans C, Boerma JT. How changes in coverage affect equity in maternal and child health interventions in 35 Countdown to 2015 countries: an analysis of national surveys. Lancet. 2012 Sep 29;380(9848):1149-56. doi: 10.1016/S0140-6736(12)61427-5. Epub 2012 Sep 20.

291 McTavish S, Moore S, Harper S, Lynch J. National female literacy, individual socio-economic status, and maternal health care use in sub-Saharan Africa. Social Science & Medicine Volume 71, Issue 11, December 2010, Pages 1958–1963

292 Robinson J, Wharrad H: Invisible Nursing: Exploring Health Outcomes at Global Level. Relationships between Infant and Under-5 Mortality Rates and the Distribution of Health Professionals, GNP per capita, and Female Literacy. Journal of Advanced Nursing 2000, 32:28-40.

293 Martin TC. Women's education and fertility: results from 26 Demographic and Health Surveys. Studies in Family Planning, 1995. 26(4):187-202

294 Ainsworth M, Beegle K,Nyamete A. The Impact of Women's Schooling on Fertility and Contraceptive Use: A Study of Fourteen Sub-Saharan African Countries. The World Bank Economic Review. 1996. 10(l):85-122.

295 Mwageni E, Masanja H, Juma Z, Momburi D, Mkilindi Y, Mbuya C, Kasale H, Reid G, de Savigny D: Socio-economic Status and Health Inequalities in Rural, Tanzania: Evidence from the Rufiji Demographic Surveillance System. Measuring Health Equity in INDEPTH Sites edition. Edited by INDEPTH Network. Accra, IDRC; 2004, Web, 2nd Jan 2015, at: http://www.equinetafrica.org/bibl/docs/MWAequity.pdf

296 Lindelow M. Sometimes more equal than others: How the choice of welfare indicator can affect the measurement of health inequalities and the incidence of public spending; CSAE Working Paper Series 2002–15; Oxford: Centre for Study of African Economies, University of Oxford. 2002.

297 Koskei EC, Koskei RC, Koske MC, Koech HK. Effect of socio-economic factors on access to improved water sources and basic sanitation in Bomet municipality, Kenya. Research Journal of Environmental and Earth Sciences. 2013, 5(12): 714-719

298 Tanzania Medical Stores Department. See here for more information: http://www.msd.or.tz/ and http://nexus.som.yale.edu/ph-tanzania/?q=node/134

299 Tanzanian Medical Stores Department. An innovative public-private partnership: Medical supply solutions in Tanzania. Yale Global Health Leadership Institute. Web, 3rd Jan 2015, at: http://nexus.som.yale.edu/ph-tanzania/?q=node/111

300 Jhpiego – An affiliate of the Johns Hopkins University. Quality Improvement of Maternal and Newborn Health Services: Introducing Standards and Recognizing Achievements. Web, 3rd Jan 2015, at: http://www.mchip.net/sites/default/files/i_Quality_Improvement_of_MNH.pdf

301 Graham W, Themmen E, Bassane B, Meda N, De Brouwere V. Evaluating skilled care at delivery in Burkina Faso: principles and practice. Trop Med & Int Health 2008; 13: 6–13.

302 Schellenberg JRMA, Adam T, Mshinda H, Masanja H, Kabadi G, Mukasa O, John T, Charles S, Nathan R, Wilczynska K, Mgalula L, Mbuya C, Mswia R, Manzi F, de Savigny D, Schellenberg D, Victora C. Effectiveness and cost of facility-based Integrated Management of Childhood Illness (IMCI) in Tanzania. Lancet 2004; 364: 1583–94

303 World Health Organisation, The partnership for maternal, newborn, and child health. A Policy Guide for Implementing Essential Interventions for Reproductive, Maternal, Newborn and Child Health (RMNCH). A multisectoral policy compendium for RMNCH. Accessed 3rd Jan 2015 at: www.who.int/pmnch/knowledge/publications/policy_compendium.pdf

304 Tanzania Nursing and Midwifery Council. Web, accessed on 3rd Jan 2015 at: www.tnmc.go.tz

305 Tanzania National Nursing Association. Web, accessed on 3rd Jan 2015 at www.tanna.or.tz

306 The Medical Council of Tanganyika. Web, accessed on 3rd Jan 2015 at http://mct.go.tz/

307 The Medical Association of Tanzania. Web accessed on 3rd Jan 2015 at http://www.mat-tz.org/

308 Africa Progress Panel. Maternal Health: Investing in the Lifeline of Healthy Societies & Economies. Policy Brief September 2010. Web, accessed on 3rd Jan 2015 at: http://www.who.int/pmnch/topics/maternal/app_maternal_health_english.pdf

309 Jowett M. Safe Motherhood Interventions in Low-income Countries: an economic justification and evidence of cost effectiveness. Health Policy. 2000 Oct; 53(3):201-28.

310 Bhandari N, Iqbal Kabir AKM, Abdus Salam M. Mainstreaming nutrition into maternal and child health programmes: scaling up of exclusive breastfeeding. Maternal & Child Nutrition. 2008. Vol 4, Issue Supplement s1, pages 5–23.

311 Bryce J, Victora CG, Habicht JP, Vaughan JP, Black RE. The Multi-Country Evaluation of the Integrated Management of Childhood Illness Strategy: Lessons for the Evaluation of Public Health Interventions. Am J Public Health. 2004 March; 94(3): 406–415. PMCID: PMC1448266

312 Mauldin WP, Ross JS, Kekovole J, Khuda B, Barkat A. Direct and Judgmental Measures of Family Planning Program Inputs. Studies in Family Planning. 1995. Vol. 26, No. 5, pp. 287-295

313 Målqvist M, Lincetto O, Du NH, Burgess C, Hoa DT. Maternal health care utilization in Viet Nam: increasing ethnic inequity. Bull World Health Organ. 2013 Apr 1;91(4):254-61. doi: 10.2471/BLT.12.112425. Epub 2013 Feb 13.

314 Sachs JD. Achieving universal health coverage in low-income settings. The Lancet. 2012, Vol 380, No. 9845, p944–947