MASENO UNIVERSITY

SCHOOL OF PUBLIC HEALTH AND COMMUNITY DEVELOPMENT

DEPARTMENT OF PUBLIC HEALTH

VOLUNTARY MEDICAL MALE CIRCUMCISION PROGRAM: EVALUATING TECHNICAL EFFICIENCY, FACTOR PRODUCTIVITY AND SERVICE QUALITY IN NYANZA, KENYA

Dickens Samuel Omondi Aduda, MPHE; MSocSc (Bioethics)

A Thesis Submitted in Partial Fulfilment of the Requirements for the Award of Degree of Doctor of Philosophy in Public Health in the School of Public Health and Community Development of Maseno University

DECLARATION

I declare that the work contained in this thesis is my original work and has not been submitted for the award of a degree at any other university. No part of this thesis may be reproduced without prior written permission of the author and/or Maseno University

Dickens Samuel Omondi Aduda, MPHE; MSocSc (Bioethics)

PG/PHD/017/2011

Signature . . ________________ Date: .12-03-2015

This thesis has been submitted for examination with our approval as University supervisors:

1. **Prof. Collins Ouma, PhD**
   Department of Biomedical Sciences and Technology,
   Maseno University, Maseno, Kenya.

Signature… … Date…12-3-15.

2. **Prof. Rosebella Onyango, PhD**
   Department of Public Health,
   Maseno University, Maseno, Kenya.

Signature……................................... Date.........………..............

3. **Prof. Jane Bertrand, PhD**
   Department of Global Health Systems and Development,
   Tulane University, New Orleans, LA, USA

Sig..……… Date.12-03-15.

DEDICATION

I dedicate this work to my dear wife, Naomi Nzingo, children Joel and Edwin and all my family members.

ACKNOWLEDGEMENT

With special thanks to:

The PhD advisory team: Prof. Collins Ouma, Prof. Rosebella Onyango and Prof. Jane Bertrand for their availability, much needed help and tremendous inputs. Many thanks to the School of Public Health and Community Development, especially its support team for their fervour, being accessible when we needed and spontaneous to help; to the Graduate School for diligent and incisive reviews of the entire work.

Special tribute to Prof. Kawango Agot, Dr. Mores Loolpapit and Mr. Mathews Onyango and the IRDO family for inspiration, contributions and many opportunities that helped to create and accomplish this work; the team that served on the Systematic Monitoring of Male Circumcision Scale up (SYMMACS) project for working so hard to support the project and collecting good data.

All my family members for cheering me on and providing the much needed support; My Christ Church Kisumu home-group members (Tom Mboya), church members and the prayer team for the prayer cover, support and encouragements.

My wife, Naomi for her willingness to spend hours-on-end in prayer support and enduring with my busy schedules while working on this project. Our children Joel and Edwin, niece Esther, Nephew Shemaiah and Celestine, a long-term support team member and friend to the family, for enduring the rigorous and demanding workload and allowing me to have space to accomplish this task.

Above all, THANK YOU God.

And thank you all.

ABSTRACT

Health service delivery across Africa is characterised by widespread inefficiency and low service quality. In Kenya, factors that closely influence performance and the critical quality measure items of medical male circumcision services remain unexplored. Specifically, technical efficiency, productivity and service quality and the factor structure of the service quality monitoring tool in Nyanza region remain unclear. Consequently, the current study aimed to evaluate performance in terms of technical efficiency, productivity and service quality of circumcision services and to explore the factor structure of quality monitoring tool in Nyanza region. Using a comparative process evaluation of voluntary medical male circumcision (VMMC) scale-up in Nyanza, site level data was collected among facilities providing VMMC in 2011 and 2012. Assessment of service tasks performed, availability of guidelines, supplies and equipment and, continuity of care was conducted using modified national VMMC monitoring instruments. Data envelopment analysis was performed to evaluate technical efficiency and productivity for 21 facilities using *PIM DEAsoft Version 3.2*. Using *SAS v. 13* software, paired t-test was performed to compare means of the obtained efficiency and productivity scores and exploratory factor analysis to clarify factor structure of quality assessment toolkit. The mean scale technical efficiency scores improved from 91% (SD 19.8) in 2011 to 99% (SD 4.0) in 2012 particularly among outreach compared to fixed service delivery facilities (CI -31.47959 – 4.698508; t= -2.8179; $df.$= 20; p= 0.005). But change in mean pure technical efficiency scores from 84% (SD 25.3) in 2011 and 89% (SD 25.1) in 2012 was not statistically significant. Benchmark facilities in 2011 were 119 and 125 but only 103 in 2012. Malmquist Productivity Index (MPI) showed service productivity declined at fixed facilities by 2.5% but gained by 4.9% at outreach ones in 2012. The indices show the improved factor productivity of 83% (p= 0.032) in 2012 was largely due to progress in technological efficiencies by 79% (p= 0.008). Principal component analysis extracted three principal factors together accounting for 29.1% of the total variance (12.9%; 9.5% and 6.7%) with final communality estimates being 13.06. Exploratory factor analysis, with item loadings ≥0.4, elicited fifteen items in factor 1, being closely related to preparedness to conduct safe procedures while factor 2 comprising five items depicts compliance with protocols in correctly performing service tasks. Using composite quality index derived from factor 1, 50% of circumcisions performed in 2011 and 58.8% in 2012 ranked as either good or excellent. The study demonstrates that facilities improved in scale but remained technically inefficient. Productivity indices showed performance was driven by technological progress from improved skills mainly among outreach facilities, but constrained by organizational and managerial factors. Facility preparedness and circumcision safety were critical service quality factors. Benchmark facilities were of fixed type. More than half of cases performed in both years ranked above average. These results provide program performance improvement objectives focussing on site level tasks, enhancing personnel technical and managerial skills, bolstering outreach services and monitoring quality using an instrument with fewer critical measure items. Further studies should explore different model estimates, effect of exogenous factors on services and routine use of composite score indices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

AIDS	Acquired Immunodeficiency Syndrome
CRS	Constant Return to Scale
DEA	Data Envelopment Analysis
DMU	Decision-making unit
HIV	Human Immunodeficiency Virus
HMIS	Health Management Information System
HTC	HIV testing and counselling
MC	Male Circumcision
MCDUs	Male Circumcision Delivery Units
MMC	Medical Male Circumcision
MoH	Ministry of Health
MPI	Malmquist Productivity Index
NACC	National AIDS Control Council
NASCOP	National AIDS and STI Control Program
SE	Scale Efficiency
SFA	Stochastic Frontier Analysis
STIs	Sexually Transmitted Infections
SYMMACS	Systematic Monitoring of Medical Male Circumcision Scale-Up
TFP	Total Factor Productivity
UNAIDS	Joint United Nations Programme on HIV/AIDS
VMMC	Voluntary Medical Male Circumcision
VRS	Variable Return to Scale
WHO	World Health Organization

DEFINITION OF OPERATIONAL TERMS

Activities: what the program does with its inputs to realise its mission.

Benchmarking: The process of comparing the performance of an individual organisation against a benchmark, or ideal level of performance. Benchmarks can be set on the basis of a sample of similar organisations; performance over time or; against some externally set standard.

Best practice: In this context refers to the set of management and work practices which results in the highest potential, or optimal quantity and combination of outputs for a given quantity and combination of inputs (*productivity*) for a group of similar organisations. Best practice can be identified at a number of levels, including organisational, national and international.

Communality: The proportion of a variable's variance explained by a factor structure. Final communality estimates are the sum of squared loadings for a variable in an orthogonal factor matrix; Communality indicates the variance in each item explained by the extracted factors; ideally, ≥0.4.

Data Envelopment Analysis (DEA): A *linear programming* technique which identifies *best practice* within a sample and measures *efficiency* based on differences between observed and '*best practice'* units. DEA is typically used to measure *technical efficiency*.

Data reduction: Reducing data to a smaller set of summary variables, with each construct measured using multiple items which can be combined in a smaller number of factor scores.

Efficiency: Degree to which the observed use of resources to produce outputs of a given quality matches the optimal use of resources to produce outputs of a given quality. This can be assessed in terms of *technical efficiency* and *allocative efficiency*.

Eigenvalue: Column sum of squared loadings for a factor. It conceptually represents that amount of variance accounted for by a factor

Exploring theoretical structure: Theoretical questions about the underlying structure of psychological phenomena can be explored and empirically tested using factor analysis

External operating environment: Factors which affect the providers of outputs that are not in the direct control of managers — for example, weather, service demand

Factor analysis: is a collection of statistical technique / methods used to examine how underlying constructs influence the responses on a number of measured variables. There are basically two types of factor analysis: (i) exploratory and confirmatory. Exploratory factor analysis (EFA) attempts to discover the nature of the constructs influencing a set of responses. It is a technique used to explore the underlying structure of a collection of observed variables; (ii) Confirmatory factor analysis (CFA) tests whether a specified set of constructs is influencing responses in a predicted way.

Factor loading: Correlation between a variable and a factor represent statistical structure of set of variables, and the key to understanding the nature of a particular factor. Respective factor loadings show the hierarchical item importance, in terms of both component availability and task performance. Squared factor loadings indicate what percentage of the variance in an original variable is explained by a factor. Correlations of items with factors (factor 'loadings')

Factor rotation: A process of adjusting the factor axes to achieve a simpler and pragmatically more meaningful factor solution **-** the goal is a simple factor structure. *Oblique factor rotation (oblimin):* Factor rotation such that the extracted factors are correlated. Rather than arbitrarily constraining the factor rotation to an orthogonal (90 degree angle), the oblique solution identifies the extent to which each of the factors are correlated. *Orthogonal factor rotation (varimax):* Factor rotation such that their axes are maintained at 90 degrees. Each factor is independent of, or orthogonal to, all other factors. The correlation between the factors is determined to be zero.

Factor: Linear combination of the original variables. Factors represent the underlying dimensions (constructs) that summarize or account for the original set of observed variables.

Factorability: the assumption that there are at least some correlations amongst the variables so that coherent factors can be identified. Basically, there should be some degree of collinearity among the variables but not an extreme degree or singularity among the variables.

Inputs: the resources dedicated to or consumed by the program.

Latent factor: A theoretical underlying factor hypothesized to influence a number of observed variables. Common factor analysis assumes latent variables are linearly related to observed variables.

Outputs: Goods and services provided to entities or persons outside the production unit.

Peers: In DEA studies, the group of best practice organisations with which a relatively inefficient organisation is compared.

Principal components Analysis (PCA): Analyses of all variance in the observed items to determine relationships with critical factors. The factors are based on the total variance. The first component accounts for the most variance in the variables. Then the second component accounts for the most variance in the variables residualised for the first component, and so on.

Production frontier: The curve plotting the minimum amount of an input (or combination of inputs) required to produce a given quantity of output (or combination of outputs).

Production technology: Relationships incorporated in production processes which determine the manner in which inputs can be converted to outputs.

Productivity: Measure of the physical output produced from the use of a given quantity of inputs. This may include all inputs and all outputs (*total factor productivity*) or a subset of inputs and outputs (*partial productivity*). Productivity varies as a result of differences in *production technology*, differences in the *technical efficiency* of the organisation, and the *external operating environment* in which production occurs.

Returns to scale: Relationship between output and inputs. Returns can be constant, increasing or decreasing depending on whether output increases in proportion to, more than or less than inputs, respectively. In the case of multiple inputs and outputs, this means how outputs change when there is an equi-proportionate change in all inputs.

Scale efficiency: The extent to which an organisation can take advantage of returns to scale by altering its size towards optimal scale (which is defined as the region in which there are constant *returns to scale* in the relationship between outputs and inputs).

Slacks: The extra amount by which an input (output) can be reduced (increased) to attain *technical efficiency* after all inputs (outputs) have been reduced (increased) in equal proportions to reach the *production frontier*. This is a feature of piece-wise linear production frontier derived when using DEA.

Technical efficiency: Conversion of physical inputs such as labour services and raw materials or semi-finished goods into *outputs*. Technical efficiency is determined by the difference between the observed ratio of combined quantities of an entity's output to input and the ratio achieved by *best practice*. It can be expressed as the potential to increase quantities of outputs from given quantities of inputs, or the potential to reduce the quantities of inputs used in producing given quantities of outputs. Technical efficiency is affected by the size of operations (*scale efficiency*) and by managerial practices (*non-scale technical efficiency*). It is defined independent of prices and costs.

Total factor productivity (TFP): Ratio of the quantity of all outputs to the quantity of all inputs. TFP can be measured by an index of the ratio of all outputs.

CHAPTER 1: INTRODUCTION

1.0 General Introduction

1.1 Background Information

Globally technical efficiency, productivity and service quality are considered critical service delivery evaluation indicators (Chisholm and Evans, 2010). The World Health Report of 2000 and 2010 (WHO, 2000) and the Africa Health Strategy: 2007 - 2015 (Chisholm and Evans, 2010; Union, 2005) underscores the strategic importance of these parameters in evaluating health systems functioning. Technical efficiency measures the ability of a unit to produce the maximum quantity of outputs for any given amount of inputs or the minimum input levels used for any given amount of outputs. Service productivity identifies 'the change in service output resulting from a unit change in the inputs' over time (Hulten, 2001). Service quality characteristics is central to enhancing the process of continuous quality improvement efforts (Hathorn *et al.*, 2011). Evaluating these service components will clarify certain aspects of medical male circumcision service delivery which remain unexplored in Kenya.

Evaluation methods may be accomplished using different approaches. Recent studies evaluating performance of voluntary medical male circumcision (Bollinger, DeCormie, & Stover 2009; Nagelkerke *et al.*, 2007; Njeuhmeli *et al.*, 2011) adopted the expenditure approach focusing on cost effectiveness analysis and potential impact, but no study has evaluated the program productivity and technical efficiency. In addition, existing evaluation studies have not considered production function in determining empirical relationships among variables of service delivery.

Voluntary Medical Male Circumcision (VMMC) program in Nyanza was rolled out in 2008. This followed recommendations by the World Health Organization (WHO) and the Joint United Nations Program on HIV/AIDS (UNAIDS) that VMMC is effective in preventing male acquisition of HIV from females. This was based on results from three randomized controlled trials conducted in Kenya, Uganda and South Africa (Auvert *et al.*, 2005; Bailey *et al.*, 2007; Gray *et al.*, 2007; Weiss, 2007) which demonstrated a 60% protective effect against HIV for men who became circumcised. As the accelerated VMMC scale-up in Kenya gets underway to rapidly cover majority of eligible clients there are legitimate concerns that this program may potentially instigate: (i) resource wastage or migration due to inefficiencies across the service provision function; (ii) heightened resource constraints as a result of diverse program components interposing on existing primary care activities and resources, thereby adversely affecting productivity; (iii) compromised service quality.

Whereas previous health system efficiency studies by Kirigia *et al*. (2002; 2004) have shown existing technical inefficiencies of service delivery at facility levels in Kenya examine the technical efficiency, productivity and service quality in VMMC service delivery in Nyanza region remains unknown. A case study by Mwandi *et al.* (2011) on Kenya’s progress in translating VMMC research into policy and program ascribed government leadership, documented implementation strategy and program flexibility as the key drivers of successful implementation of scaled-up program since 2008. While the study provides a global picture on program performance, it does not explore factors of service delivery function that are more closely related to its productivity and efficiency. Evaluation of these dimensions is necessary to clarify performance of VMMC service delivery in various contexts and provide critical information for management decisions.

Facility assessments of VMMC service delivery conducted in Nyanza region soon after the program roll-out revealed existing widespread systems and facility level inadequacies, in terms of working space, training and availability of skilled staff (Herman-Roloff *et al.,* 2011). Such constraints can impact adversely on production of VMMC services. Evaluation studies conducted by: *(i)* Hankins *et al*. (2011) to review cost, impact, and challenges of accelerated program scaling up and, *(ii)* Bertrand *et al*. (2014) to evaluate progress of VMMC and monitor elements of surgical efficiency did not elicit factors closely related to variations in technical efficiency, productivity and service quality dimensions of the program scale-up. Furthermore, no benchmarking methods for service delivery facilities in Nyanza has been considered or examined.

Specific studies (Bertrand, et al., 2014; Rech, et al., 2014a) conducted on determinants of surgical efficiency of VMMC service delivery and anecdotal field reports have specifically highlighted existing organizational and managerial challenges including weak adoption of production technology and lapses in compliance with guidelines. In these studies, surgical efficiency was conceptualized based on six elements recommended by WHO: use of multiple surgical bays, use of kits containing pre-bundled consumables and disposable instruments, task-shifting, task-sharing, use of the forceps-guided surgical method, and use of electrocautery (diathermy) to stop bleeding more quickly than suturing (WHO, 2010a). Surgical efficiency outcomes were defined in terms of the primary provider's time with the client (PPTC) and total elapsed operating time (TEOT). While these studies determined that improved TEOT did not lead to poor service quality, critical quality measure items as well as specific components of technical efficiency and productivity variation in VMMC service delivery in Nyanza remains unknown.

Recent reports (Jennings *et al.,* 2014; MOPHS, 2011) have highlighted global concerns about the capacity of 'health systems to deliver and sustain VMMC according to minimum quality criteria'. The study by Jennings *et al.* (2014) examined VMMC program performance from a health systems perspective in terms of compliance with quality standards. It observed that service quality was related to facility preparedness to provide safe VMMC services, availability of effective support supervision, documentation and reporting and pressure on staff. However, the factor structure of the existing VMMC quality monitoring tool and the items critical for service quality measure are still unknown. This is necessary to improve effectiveness of site level quality assessments and ranking VMMC service quality performance. Poor service quality can jeopardise realisation of both the program's intended direct HIV intervention goals as well as the ancillary objectives such as capacity building (training of personnel as well as provision of equipment and supplies) meant to augment strengthening of ambulatory health care particularly at the peripheral health service levels (UNAIDS, 2009; WHO/UNAIDS, 2008).

The National AIDS & STI Control Program (NASCOP) is currently seeking to establish national mechanisms for objective program assessments and improved reporting to ensure effective ongoing improvements in service production and quality management plans based on a minimum criteria (National AIDS & STI Control Program, 2009). It is thus imperative to evaluate the multiple dimensions of VMMC service delivery to identify changes associated with different aspects of production factors. This study employs data envelopment analysis and factor analysis techniques to evaluate technical efficiency and productivity of VMMC service delivery data as well as service quality since they can handle multi-dimensional service delivery data and summarize results as a score index which is simple to interpret. The results will contribute useful evidence for enhancing decision making in development of guidelines and monitoring tools.

1.2 Problem Statement

The clinical efficacy of male circumcision has been demonstrated through a series of observational and randomized controlled studies. However, globally, there is no clear understanding of, nor a consensus on the most efficient models for service delivery that would optimize performance within resource constrained contexts. However, which types of service delivery are likely to be efficient in terms of optimal performance given a set of input-output variable mix in Nyanza Province remain unknown. Furthermore, there are concerns on how best to aggregate and clarify complex program data and benchmark VMMC service producing units.

Existing needs assessment reports, operations research results and program data consistently show widespread systems and facility level inadequacies, in terms of constrains in working space, equipment and supplies management, inadequate or lack of compliance to standards, potential for widespread practice variations and staff fatigue or burnout. No studies exist to clarify critical implementation behaviours and performance of service units in terms of technical efficiency, factor productivity and service quality. The existing quality assessment tool is reportedly too laborious for service providers to use routinely, and the information collected complex making its use problematic. In addition, there are concerns related to aggregation, interpretation and reporting multidimensional data on VMMC service delivery. Consequently, the current study uses non-parametric data envelopment analysis (DEA) which incorporate quality variable to evaluate efficiency and productivity. Factor analysis is also used to provide an objective method to elicit key quality measure items of MC service delivery.

1.3 Study Justification

In Kenyan health sector, efficiency monitoring and quality assurance and standards are program policy priority issues. This is in line with a policy initiative adopted in 2006 by the African Union for member states to institutionalize efficiency monitoring within their health sectors to enhance management oversight decisions. Evaluating VMMC service based on program delivery functions provides baseline information on critical aspects of program efficiency, quality and productivity by examining the scope of productivity performance, sources of process compliance and constraints. The study objectives to assess specific indicators of technical efficiency, bench-marking and simplifying the factor structure of the currently recommended quality toolkit for monitoring service delivery at MC facilities provide critical information for improving on reporting program data and decision-making.

The study findings will guide policy makers and implementing institutions to focus efforts in the most productive area of the operational process in designing effective improvement strategies. Additionally, the identified latent factors underlying the existing service quality assessment toolkit and their respective item coefficients will inform development of simpler but sensitive quality assessment tool.

1.4 Objectives

1.4.1 General Objective.

To evaluate the technical efficiency, productivity and service quality in delivery of voluntary medical male circumcision and to explore the factor structure of the service quality monitoring tool in Nyanza region during 2011/2012.

1.4.2 Specific Objectives.

1. To determine technical efficiency for sampled male circumcision facilities in Nyanza region during 2011/2012 using data envelopment analysis.
2. To identify the benchmark peers for observed VMMC facilities in Nyanza region based on the DEA model.
3. To characterize productivity changes in VMMC service delivery in Nyanza during 2011-2012 using DEA Malmquist productivity index.
4. To identify latent factor dimensions and item interrelationships of the VMMC service quality assurance tool using factor analysis techniques.
5. To rank service quality performance of circumcisions in Nyanza region using quality index comprising the critical items of the quality toolkit.

1.5 Research Questions

1. What is the technical efficiency scores by data envelopment analysis for sampled male circumcision facilities observed in Nyanza region in 2011/2012?
2. Which are the benchmark peers units for the observed VMMC facilities in Nyanza based on the DEA model?
3. What are the characteristics of variations in productivity of VMMC services in Nyanza using DEA Malmquist productivity index?
4. What are the latent factor dimensions and item relationships of the VMMC service quality assessment tool based on factor analysis techniques?
5. What are the quality rank performances of circumcisions in Nyanza region based on composite factor index derived from the main latent factor of the quality assessment tool?

CHAPTER 2: LITERATURE REVIEW

This chapter is organised by objectives. Literature related to the basic theories, concepts and measurement approaches of program efficiency, factor productivity and service quality assessments is reviewed. The context of the study is scaled-up VMMC program in Kenya. Section 2.1.0 highlights the basic study concepts, key definitions and evaluation frameworks. Sections 2.2.0 – 2.7.0 focus on reviews related to specific study objectives.

2.1.0 Background Information: Study Concepts and Evaluation Frameworks

2.1.1 Implementation of VMMC in Kenya

Voluntary medical male circumcision (VMMC) for HIV intervention was rolled out in Kenya in 2008. Evaluating its service delivery strategies and processes provides insight into how well they work to achieve desired goals in terms of outputs produced vis-à-vis resources used. It also provides information on the technical and functional capacity of production units/process to accomplish standard service delivery tasks (Betrand, et al.,2014; Cooper, Seinford & Tone, 2007; Herman-Roloff et al., 2011; Mwandi et al., 2011).

As a program VMMC service delivery *(i)* is characterized by a multi-component resource intensive function which has considerable implications on its quality and that of other on-going health intervention services (Chen *et al.,* 2004; WHO, 2010a); *(ii)* efficiency and productivity portends the program's impact on HIV epidemic and policy directions (Wegbreit *et al.,* 2006); *(iii)* coverage required to realize the intended public health impact is large and expansion rate rapid (Kahn *et al.,* 2006; WHO, 2010a); *(v)* quality and resource use improvement objectives hinge on institutional and micro-level organizational factors (Chen *et al.,* 2004). Improvement objectives may focus on: healthworker performance; environmental attributes of the workplace; behaviour and / or receptiveness of client and; provision of space, tools, equipment, and supplies to enable health workers to do their jobs (Derose *et al.,* 2002). While implementation variation is inevitable in scaled-up public health interventions, it is necessary to assess how organizational factors contributes to intervention outcomes and be able to recognize and define key performance improvement needs. Since there is still paucity of information on performance of VMMC service delivery production function in Kenya, the current study evaluated technical efficiency, productivity and service quality of VMMC services in Nyanza region. Data envelopment analysis and factor analysis were used to summarize the complex multidimensional data.

2.1.2 Overview of service delivery assessment.

Service delivery is the key function of health systems blocks. It is defined as 'the way inputs are combined to allow provision of a series of interventions or health actions' to promote, restore or maintain health in an equitable manner (WHO, 2000). Various concepts exist for evaluating service delivery. The prevalent perspectives consider: *(i)* the relationship between inputs (such as manpower and capital) available for service delivery and the outputs (including services, products, or technologies) that results from health care activities (this is known as productivity perspective); and *(ii)* performance of service delivery in terms of the health effects or status change resulting from the outputs (effectiveness perspective) (Veihola, 2008).

Recent studies evaluating performance of voluntary medical male circumcision (Bollinger *et al.,* 2009; Nagelkerke *et al.,* 2007; Njeuhmeli *et al.,* 2011) adopted the expenditure approach focusing on cost effectiveness analysis and potential impact, but no study has evaluated VMMC program productivity and technical efficiency. In addition, existing evaluation studies have not considered production function in determining empirical relationships among factors of service delivery.

The present study focused on evaluating performance of VMMC service delivery in Nyanza, Kenya as a public health (in contrast to personal health) intervention. The interest in adopting a service delivery production function is to facilitate assessment of effects of shifts in service input-output mix on the marginal rate of service productivity for VMMC. Marginal rate has been defined as 'the change in service output resulting from a unit change in the delivery inputs' (Murray and Frenk, 1999). Productivity and technical efficiency measure the ability of a unit to produce the maximum quantity of outputs for any given amount of inputs or the minimum input levels used for any given amount of outputs. These measures are useful when considering rationalizing service activities to enhance production by focusing efforts in the most productive area of the operational process (Breen *et al.,* 2002). In Nyanza region, the technical efficiency and key drivers of productivity of VMMC service delivery is still unknown. The current study evaluated technical efficiency and productivity of facilities to establish their relative resource use vis-à-vis outputs and identified benchmarking VMMC facility units.

To evaluate technical efficiency and productivity of VMMC services in Nyanza region, the current study adopts the productivity perspective focusing on resource mix rather than effectiveness since change in health status from male circumcision activities occur in the long-term. It is difficult to attribute these long-term outputs to the input period assessed (Yanshuang and Byungho, 2012).

Globally, evaluating performance of public health intervention programs is often encumbered with challenges such as conceptualizing comprehensive measures; designing appropriate evaluation models; harnessing composite data and reporting assessment information; attributing variability to specific program measures and; applying the results to improve program goals (Derose, *et al.*, 2002). The current study applied data envelopment analysis and factor analysis tools since they are amenable to handling multi-dimensional program data (both inputs and outputs) simultaneously and summarizing the results in terms of composite index scores, which are simpler to interpret. The methods have not been applied before in analysing VMMC data in Kenya. They enhance opportunities to uncover relationships that remain hidden for other methodologies (Akazili *et al.*, 2008). As such the study results will provide critical information for objectively identifying performance progress, improvement needs and potential intervention efforts (WHO, 2010a, b) for VMMC services in Nyanza region.

2.1.3. Evaluation framework.

Evaluating dimensions of service delivery function may focus on the structure, process, intermediate output, and outcome or impact levels (Fig. 1) (Hollingsworth, 2008; Issel, 2004).

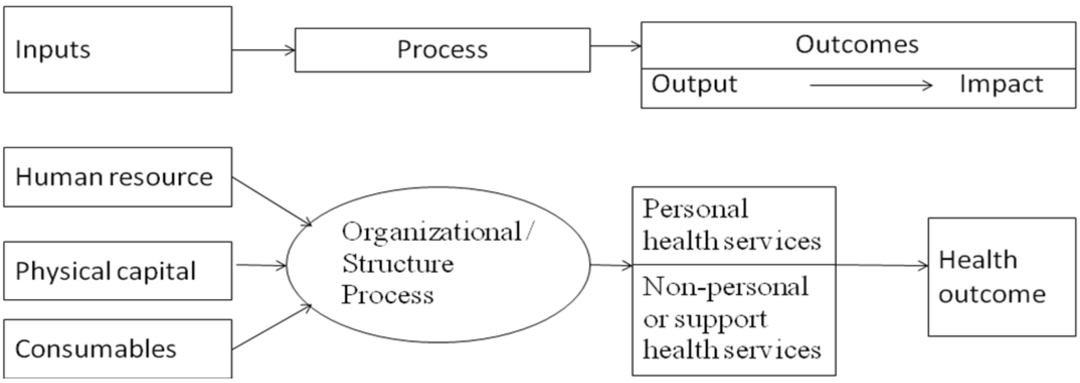

Fig. 1: A framework for evaluating the service provision function focusing on health system inputs, processes and outputs (WHO, 2002)

Specific component variables include (as used in the study): *i) inputs* (clinicians, nurses, surgical bed, surgical time); *ii) process* (organizational/structure such as service delivery) and; *iii) outcomes* (output – services including number of circumcisions accomplished, proportion of circumcised men receiving HIV test, service quality; impact – ultimate health outcome desired – not considered here). The users'/providers' perspectives are considered an extension of the outcome dimension (Adams *et al.*, 2003).

While measurements may be performed on specific or multiple dimensions at different levels depending on the evaluation goals (Hollingsworth, Dawson, & Maniadakis, 1999) the current study considered all activity dimensions of the circumcision procedure. This is more preferred for comprehensiveness (Hollingsworth, 2008; Lukas and Hall, 2010). Additionally, to improve choice of service dimensions to evaluate, it was necessary to consider the appropriateness of the services or procedures performed, the resources used and whether the outcomes were desirable or worthwhile (Leah and Barbara, 2009). Performance evaluation of circumcision tasks focusing on technical efficiency, productivity and quality was preferred to enhance homogeneity.

The direction and magnitude of performance parameter variations in delivery of the chosen intervention reflect changes in its service functions, depending on the organization, fiscal factors, systems structures and process (Reinhardt and Tsung-mei, 2000; WHO, 2004). An efficient service delivery unit is one which is able to produce 'optimal' quantities of output within the available resources (input quantities) given the type of technology (process) feasible at a point in time. The meaning of "optimal" defines the meaning of "efficiency" among peer or homogeneous group of health facilities under comparison (WHO, 2010b).

The national VMMC program is a new HIV intervention technology that still being implemented as a vertical program and is resource intensive. As such, it is possible that implementation across diverse contexts may be affected disproportionately by differential funding, resource deployment models and operational issues. In Nyanza, the magnitude and direction of resource use by VMMC facilities have not been defined, hence the need to evaluate technical efficiency and productivity of circumcision services.

2.2.0 Technical Efficiency for Sampled Male Circumcision Facilities in Nyanza Region Using Data Envelopment Analysis

2.2.1 Background information from previous studies.

Data envelopment analysis to determine technical efficiency and productivity measures of health system in Africa (Alexander, Busch, & Stringer, 2003; Chisholm & Evans, 2010; Kirigia, Emrouznejad, Sambo, Munguti, & Liambila, 2004; Torkian and Ranjbar, 2013) have shown that adverse variations in the parameters may be due to inappropriate or costly staff mix; underutilized staff and equipment; irrational use of consumables (drugs and other supplies); sub-optimal quality care; inappropriate facility size, corruption and fraud largely due to poor governance; insufficient and/or irrational budget and resource allocation. Specifically in Nyanza region, the few studies that have evaluated VMMC program performance only provide global information about its successes and challenges (Mwandi, *et al.*, 2011); social and functional aspects as well as lessons learned during implementation (Herman-Roloff *et al.*, 2011); adverse events (Bailey *et al.*, 2008); conducting accelerated service delivery activities (National AIDS/STI Control Programme, October 2010); program coverage needs and logistics (Nagelkerke *et al.*, 2007) and monitoring of activity performance and compliance with minimum criteria (Bertrand et al., 2014; Obiero et al., 2011). However, technical efficiency, productivity and identification of key quality characteristics of VMMC service delivery in Nyanza region have not been explored. Consequently, the current study evaluates these dimensions using data envelopment analysis (DEA) and exploratory factor analysis.

Previous DEA studies in Kenya (Jennings *et al.,* 2014; Kirigia *et al.,* 2002) to assess technical efficiency of health care focused on hospitals and primary health care program but did not evaluate productivity change and did not include quality dimensions in their DEA models. The present study therefore considers these aspects by conducting a comparative process evaluation of sampled sites providing VMMC in Nyanza.

2.2.2 Technical efficiency of facilities providing VMMC services.

2.2.2.1 Conceptual considerations.

Technical efficiency of health service delivery is essentially concerned with how tasks and interventions are applied in 'attaining the highest level of health possible within the available resources'(WHO, 2004). It is considered a key measure of service delivery function outcomes (Union, 2005; WHO, 2000, 2010b; WHO/UNAIDS, 2007). It is defined as 'the maximum amount of health service outputs and best outcomes obtainable over a period of time from a given set of input combinations within a given setting' (WHO, 2000).

In the more intermediate terms, it refers to using fewer (same) input amounts or combinations to get the same (better) health outputs from service delivery while maintaining quality of care (Chisholm and Evans, 2010). Theoretically, this definition presupposes a production function, which in turn assumes the process is resource/cost minimizing while maximizing productivity. Technical efficiency (global TE) is a product of pure technical (PTE) and scale efficiency (SE). Pure TE is generally associated with organization of operations of the specific service producing units (input-output mix) while scale depicts issues related to size. Sources of inefficiency of a facility unit may thus be attributable to either or both of the components (Cooper *et al.*, 2007; Murray and Frenk, 1999).

2.2.2.2 Analysis of technical efficiency.

Assessment of technical efficiency may be accomplished using ‘best performance frontier’ or ‘central tendency or average-based’ techniques. The two perspectives can potentially result into different improvement decisions (Cooper *et al.,* 2007). Furthermore, there exist variants of either type of analyses. Whether to prefer either one or a combination of the methods depends on the study context and objectives, data characteristics and user skills.

The frontier methods include non-parametric data envelopment analysis (DEA) and parametric stochastic frontier analysis (SFA). Both can be used to identify a production frontier for a group of facilities but they employ different assumptions and methodologies. DEA methods use mathematical programing to obtain the production frontier enveloping all the observed data unlike SFA. It measure the extent to which outputs can be increased through higher efficiency without using additional resources (inputs) under different scale size assumptions. It also identifies efficiency into its technical, allocative and economic components (Burgess, 2012b).

Limitations of DEA include sensitivity to outliers, assumes no errors (which may bias results) and standard models do not permit hypothesis testing for the best model specification (Lee and Holland, 2001), Conversely, stochastic frontier methods are parametric, typically accommodate only a single input with multiple outputs; differentiates errors from inefficiency sources; requires specification of a functional form and; permits computation of the confidence intervals for efficiency scores and their best predictors for individual facilities. However, based on parameter estimates it may not envelop all output points and does not identify peers (Lee and Holland, 2001). Regarding regression methods least squares are used to define functional relationships between one dependent variable and other or multiple independent ones and predict sources of variations. The methods estimate a single sample-based global average score and are amenable to hypothesis testing.

Increasingly, data envelopment analysis (DEA) is becoming instrumental in evaluating health service delivery efficiency. This may be attributed to: *(i)* its capacity to handle multiple inputs and outputs of any measurement (both controllable and those beyond a provider's control) and dimension (Cooper, *et al.*, 2007) to produce a single aggregate relative "efficiency score" for each unit. These scores, adjusted to be a number between 0 and 1 (0 – 100%) are relative measures estimated based on the most favorable combination mix for each unit in contrast to using an absolute standard; *(ii)* DEA incorporates all the data available to construct a 'best practice' frontier and simultaneously compares facilities to classify each unit most favorably; *(iii)* ability to identify respective unit productivity individually, sources of inefficiencies as well as the benchmark members ('peers'= units assigned a score of 100%) in the set and the scope of improvements required for the less efficient ones, though it does not reveal how to accomplish the needed changes. Ideally, the changes selected by management for respective facilities should consider their practicality and feasibility (Hollingsworth, 2003; Naveen and Boonghee, 1998) and; *(iv)* it does not require inclusion of cost variable nor modelling of functional relationships for inputs to outputs (Zhu, 2003).

Specifically DEA estimates efficiency scores for each unit (a composite score aggregating values of all included model variables, regardless of unit or dimension) by comparing its input mix and volume of services provided against the best performing peers in the set (Worthington, 2004), depending on the preferred model orientation and service delivery unit size. Model orientation provides information on how efficiency scores are derived and how they vary. Technical efficiency indicators may be either input-orientated or output-orientated depending on which variable set the program managers have control. Input-orientated technical efficiency focuses on minimizing inputs used without reducing the output quantity while in output-orientated efficiency the focus is on expanding output quantities while maintaining current level of inputs. The present study assumed an output orientated evaluation (2007).

Additionally, it is necessary to indicate if the facility is too large or too small considering the inputs used to produce the observed outputs. This is accomplished by examining scale efficiency under different model versions which make different assumptions about returns to scale: constant returns to scale (CRS), variable returns to scale (VRS) and non-increasing returns to scale (NIRS) (Worthington, 2004). Variable return to scales (VRS) assumption allows the frontier level of outputs to inputs to vary with the size of the facilities assessed and unit comparison is restricted to only among those with comparable sizes whereas under constant return to scale (CRS) the frontier (‘best practice’ level) is determined by the highest achievable ratio of outputs to inputs for each unit, regardless of size.

The technical efficiency scores estimated are affected by the model characteristics, number of model variables and level of input variables used by best performing facilities in terms of their outputs to inputs (normally the resources needed to compete a task) ratio. The scores reflect the performance of each facility relative to best performing ones. The exact interpretation depends on the DEA model orientation used: whether output-maximizing or input-minimizing (Hollingsworth, 2004). The obtained scores are considered consistent estimators of the true (unobserved) efficiency of the service units. The position in relation to the production frontier may be due to movements in efficiency factors (contextual variables), the inefficiency of production units, or random noise (Banker, 1993; Hollingsworth, 2006).

Traditionally, performance of programs have been assessed against absolute standards estimated as global average values, mainly focusing on controllable input variables such as staff and capital. The strategic importance of comparative performance evaluation of VMMC service outlets is that each unit is ranked according to the most favourable performance relative to similar ones in the set within given contexts. The DEA outputs provide diagnostic performance information for a set of comparable service delivery units. Nevertheless, it is desirable that management decisions also consider broader policy objectives such as service access and coverage as well as other exogenous factors. In the current study, extraneous factors were held fixed in the DEA model used, being outside the purview of providers.

2.3 Identifying Peers and Benchmark Units for the Observed VMMC Facilities Based on the DEA Model

The 'Kenya Health System Assessment Report (2010), indicates the import of obtaining information on efficiency of service delivery to inform benchmarking of health system improvement activities. Using DEA for benchmarking service units is useful since it incorporates multiple variables regardless of units of measurements. Benchmarking is practically an 'accounting process' and assumes that service units are engaged in a production process which transforms inputs into outputs. It can be used in identifying 'best-practice' (reference or peer) production units to guide performance targeting and planning improvement activities. The peer units demonstrate optimal production models by providing a 'standard of excellence' or 'best practice' reference for the corresponding inefficient ones in terms of resource types and combinations. The DEA techniques estimate potential variations required to make each inefficient unit as efficient as the most efficient (best-practice) ones on the frontier or best practice level (Zhu, 2003). This information can be used by decision-makers as a means of projecting improvement needs for each service production unit in respect of the peer.

While benchmarking can be accomplished by using average-based methods such as ordinary least squares (OLS), the frontier-based DEA is preferred. The average-based methods use average level of performance to compare production units in the set, whereas frontier-oriented techniques compare performance based on the efficient frontier or best practice (Zhu, 2003). In DEA-benchmarking, as used in this study, multiple performance measures (inputs and outputs) are included in a single integrated model to generate comprehensive information on 'best-practice' units. Depending on the purpose of the evaluation, the DEA-benchmarking provides different sets of comparison peer groups. These include efficient units ('best practice') peers plus the option production combinations for the respective less efficient service units. This makes it possible for program managers to identify potential efficiency improvements (whether to reduce or increase production factors) to reach efficiency targets.

In an output orientation the decision question regards ‘how much can the inputs be cut while increasing the outputs’? On the other hand, in an input orientation, the question is ‘how much can the outputs be increased while reducing the inputs? After identifying the target improvement values, there may still exist ‘slacks’ to deal with. Slacks indicates the scope for improving input and output values after the changes in input and output levels corresponding to the optimal value have been netted out. A service production unit is truly efficient if it has no slacks. Since technical efficiency has not been defined before for VMMC service delivery in Nyanza to show the ‘optimal’ performance of facilities, the present study performed a DEA-based benchmarking to define ‘best practice’ peers among the sampled facilities.

2.4 Characterizing Changes in Productivity in VMMC Service Delivery Function Based on DEA Malmquist Productivity Index

Facilities providing VMMC services deploy resources with varying levels of efficiency. Productivity of service delivery is determined by differences in production technology, technical efficiency of the organization, and the external operating environment in which production occurs over time. While studies have provided evidence on the rationale for VMMC (Kahn *et al.,* 2006; Nagelkerke *et al.,* 2007) and related implementation experiences (Herman-Roloff *et al.,* 2011) objective evaluation of productivity and related factors have not been reported. It is still unclear how changes in marginal service productivity for VMMC, defined as ‘the change in service output resulting from a unit change in the inputs’ (Hulten *et al.,* 2001) vary across facilities over time, in diverse service provision contexts and what factors drive the changes. Assessing the extent to which VMMC facilities differ in their resource use to accomplish service delivery tasks shows changes in their respective production in relation to the best performing units in the set from one time point to the next as well as factors closely associated with these changes.

The current study postulates that efficiency and productivity of the medical male circumcision scale up in Kenya is sensitive to the factors that providers have control over (discretionary), such as the number of staff and technology factors (WHO, 2010a; World Health Organization, 2007) as well as those beyond providers' purview (exogenous/non-discretionary) including the epidemiological context of the disease and distribution of population behaviours (Wegbreit, *et al.,* 2006). Following Banker and Morey, 1986 (Banker and Morey, 1986) the exogenous/non-discretionary factors are considered fixed (since they are out of the control of the providers) so that while they are allowed to influence the relationship between inputs and outputs, the "efficiency score" is not affected by them.

Assessing resource use and productivity require consideration of multiple dimensions including inputs used, outputs generated and service quality (Arah *et al.,* 2003) to provide a platform for demonstrating specific sources and amounts of variations in the service delivery process and outcomes. Assessment indicators would normally be designed in relation to one or multiple aspects of these dimensions (Burgess, 2012a; Lukas and Hall, 2010). A composite measure (which is a combined metric that incorporates multiple individual measures across the production spectrum to provide a single score) (Fernando and Abeynayak, 2012) is normally preferred to simplify evaluation of multiple dimensions. Combining the critical individual service delivery components into a single metric enhances efforts to improve the underlying aspects supporting the provision of services by the facilities (WHO, 2010a; WHO/UNAIDS, 2009). The results generated can be used to compare target service units (Murray and

Frenk, 1999) across different delivery dimensions and activity levels (Smith & Andrew, 2004; WHO, 2004).

Malmquist productivity index (MPI) – defined as output per unit of input resources used – is a process-oriented composite measure suitable for the multi-level and multi-dimensions health care performance data (McGlynn *et al.,* 2008). It measures relative shifts in productivity over time for each production unit based on the prevailing/common technology . The MPI specifically identifies whether the observed productivity changes (shift in production) involve pure technical, scale, or technological efficiency change (Hollingsworth, 2008). For an output-based productivity index, values greater than 1 indicate improved or positive growth (progress) from base period, while values less than 1 show negative growth (regress or decline) in productivity elements (Chowdhury *et al.*, 2010). Variations in factor productivity measures, estimated as a geometric mean, can be accounted for in terms of technical change and efficiency change (Kounetas and Tsekouras, 2007).

The technical change indicates a shift in technology in the best-practice frontier itself between period $t + 1$ and t. Improvements in technology and/or size will produce an upward shift effect in the production frontier itself. Pure technical inefficiency indicates poor management of labour/capital inputs and the score usually reflects the amounts by which efficiency should be improved. However, it also reveals the differences which exist between different contexts (Steering Committee for the Review of Commonwealth / State Service Provision, 1997). Progress in scale efficiency is largely attributable to better organization with increase in size of the unit (Chowdhury *et al.,* 2010; Färe *et al.,* 1997). Organizational changes that lead to improved use of resources will move the service unit position towards the ‘best practice’ frontier (a catching up effect) indicating positive efficiency gains. If the “catch-up” is larger than “frontier shift”, the changes are largely due to efficiency attributable to improved organizational management (Kounetas and Tsekouras, 2007).

Technical change in VMMC program may result from new inputs or service quality improvements over time in existing program’s technology input factors, including technical procedures, equipment and labour-force skills. Changes might also be occasioned by change in operational structures (Eklund, 2008; Hulten *et al.,* 2001) for example when implementing partners restructure to improve organizational functions and efficiency. Such changes may emphasise utilization of treatment protocols and referrals; support in terms of supervision, logistics and training; and management which includes planning, budgeting monitoring and evaluation. The organizational structure of the agencies also determines the working environment for staff, their motivations, clarity of organizational goals, decision-making, staff mix and work responsibilities (WHO, 2010b). In addition it influences the distribution of the quantity and quality of health services that can be provided to the population in relation to their health care needs (NASCOP, 2009). Whereas VMMC service delivery involves substantial resources across diverse and complex contexts, it is still unknown how productivity parameters change over time in these settings.

2.5 Determination of Latent Factors of Service Quality Assurance Tool Using Exploratory Factor Analysis

Although a precise definition of service quality remains elusive (Parasuraman, Zeithaml, & Berry 1985) there is consensus about its multidimensionality as a service production variable (Hathorn *et al.,* 2011). The Institute of Medicine (IOM) defines quality as “the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (Derose, 2002). Inherent in this definition are dimensions related both to health experiences and anticipated program activity outcomes. The quality dimensions mentioned in literature vary but they largely encompass the structure, process and outcome aspects. Guidelines for program implementation ought to address how these dimensions can be correctly assessed across diverse care settings. Accordingly valid data tools are needed to capture objective service quality information during routine program practice. This would enhance operational decision-making (WHO, 2009); determine scope for resource allocation and improvement tasks (World Health Organization, 2006); enhance accountability for service delivery tasks planned or accomplished; guideline revisions (Derose and Petitti, 2003; Rapkin *et al.*, 2008) and facility accreditation (Belows & Sullivan, 2004). Furthermore, the structure of data tools should be designed to improve their effectiveness (Kredo et al., 2012; Lazar et al., 2013). The current quality assessment tool for VMMC services is laborious to use and critical measure items are still unclear.

Voluntary medical male circumcision (VMMC) service delivery occurs across multiple service levels in diverse contexts which pose considerable potential for variability in service quality (Herman-Roloff et al., 2012; Mahler et al., 2011; Mangham & Kara, 2010; MOPHS, 2011; NASCOP, 2008a, b, 2009). Specific issues relate to lack of adherence to guidelines; (Campbell *et al.,* 2000; Nietert *et al.,* 2007) low level support supervision; constrained documentation (Bertrand et al., 2014; Hong et al., 2006; Jennings et al., 2014), reporting and uptake of feedback (Mahle *et al.,* 2011; MOPHS, 2011).

These have implications on service uptake, safety and improvement decisions. Furthermore, hurdles in service quality assessment occur in relation to conceptualizing comprehensive measures (Hong *et al.*, 2006); conducting appropriate evaluations; aggregating and reporting assessment information; attributing variability to specific service quality measures and (Shaller, 2004); applying the results to improve program goals (Bellows, & Sullivan, 2004).

Considering these challenges and as part of efforts to ensure sustained quality of services, the national VMMC program adapted for use in 2009 (MoH, 2005; NACC, 2009; NASCOP, 2009) the World Health Organization (WHO) toolkit (WHO, 2009) for monitoring a range of quality standards at facility level. The instrument is a comprehensive checklist comprising of 10 standards and 36 criteria. Additionally, the Kenya Quality Model for Health (KQMH) which was launched in 2012 provides a broad framework for sectoral integration of service quality improvement and management. This study explored the status of VMMC service quality and assessment instrument.

While this is a useful instrument for internal and external assessment of VMMC service quality activities across multiple levels, it is reportedly complicated and laborious. Furthermore, there is need to determine constructs that relate closely to the local context since quality presentations may be influenced in part by cultural factors (Kredo *et al.*, 2012).

Using exploratory factor analysis technique (a procedure employed to reduce the number of variables based on their intercorelations, while retaining most of the observed information and to determine if the original variables are organised in a particular pattern reflecting another latent variable) the current study explored the underlying dimensions and variable interrelationships of the items comprising the SYMMACS service quality assessment tool (adopted from the WHO quality toolkit (Floyd and Widaman, 1995; WHO, 2009); identified key constructs, which demonstrate optimal service quality performance measurement (principal component analysis technique was used to extract the key components explaining the greatest proportion of the observed variation in the data); and ranked the study sites based on the quality factor scores (quality index) derived from the main component extracted according to their respecting factor loadings. The study thus fills the gap by identifying critical quality characteristics and simplifying the toolkit by clarifying essential item measures. These can be adopted by managers for routine service quality assessment to identify sources of variations in service outcomes.

Historically, service quality has been assessed via similar basic dimensions (Parasuraman, *et al.,* 1985; Zaneta and Ilona, 2008). The Donabedian framework, (Donabedian, 1988) for example, proposes 'structure – process – outcome' approach. In this approach, the structure dimension relates to the care setting (including facility characteristics, equipment, training and special skills). The process dimension comprises aspects related to the provider-patient interaction and is considered a function of the technical and interpersonal skills. The outcome dimension reflects the immediate/intermediate and long-term changes occurring to the patient's status based on services provided (Derose *et al.,* 2002; Mangione-Smith *et al.,* 2007). The current study considered multiple dimension approach in collecting service quality data. The critical item measures of VMMC service quality are still unknown. Based on the study data, the critical quality measure items were identified and used to construct a composite quality index for evaluating circumcisions performed across VMMC facilities in different settings.

CHAPTER 3: METHODOLOGY

3.1 The study area

Nyanza is located in the western part of Kenya, at the shores of Lake Victoria. It covers a total area of about 12,477.1 km2 (4,817.4 sq. mile). The total population in 2009 (2009 Census) was estimated at 5,442,711. It is divided into 6 Counties. The major economic activities in the region include farming, fishing and other business enterprises (Figure 3.1.).

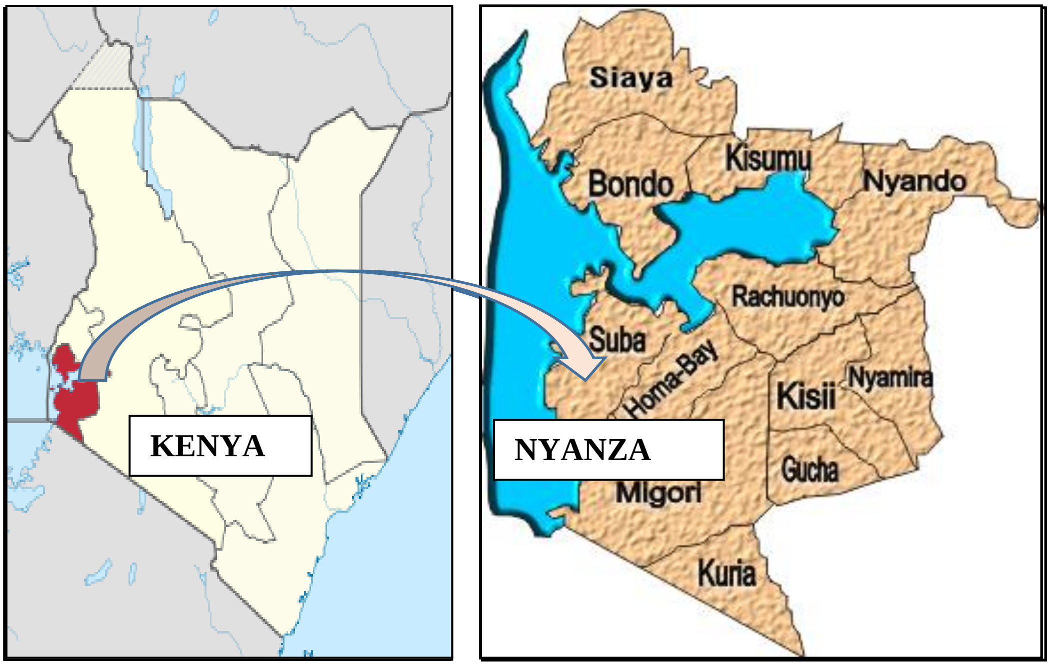

Fig. 3-1. Study Area: Nyanza Province ***(Source: a Google map download)***
Location in Kenya: Coordinates: 0°30′S 34°40′E

The present study was as a sub-study within the SYMMACS Project conducted in the VMMC program areas in southern Nyanza – Migori, Homa Bay, Ndhiwa, Rongo, Nyatike, Suba, Mbita districts; central part of Nyanza – Nyando, Kisumu East and Kisumu West and western part of Nyanza (Siaya, Ugenya, Gem, Bondo, Rarieda.Nyanza region. Being the focus of the mop-up phase of the VMMC national program, Nyanza hosted the largest VMMC implementation programs with the greatest concentration of resources, although this varied with the stage of implementation.

3.2 Epidemiological Information

In 2009, the overall HIV prevalence in Nyanza was 13.9%, but varied between various sub-populations at high risk. Adult HIV prevalence across areas occupied by traditionally non-circumcising communities (*Luo-Nyanza*) ranged between 16.4 in Migori to 22.5% in Nyando compared to 4.7% and 7% among the neighboring *Kisii* and *Luhya* communities respectively and 6.3% nationally (KNBS, 2010). The Nyanza HIV epidemic is mainly driven by lack of male circumcision (Bailey *et al.,* 2007; WHO/UNAIDS, 2007) the cultural practice of widow inheritance as well as polygyny and other forms of concurrent sexual partnerships (Gray *et al.*, 2000). By 2010, there were 332 health facilities in Nyanza and 260 of these were current or had previously supported VMMC services.

3.3 Study Design

A comparative process evaluation of Voluntary Medical Male Circumcision Scale-Up in Kenya was conducted over two years, targeting the VMMC facilities and tasks accomplished during circumcision procedures (the units of analysis). Site level data was collected among facilities providing VMMC services (categorized as fixed, outreach or mobile) using a modified national quality monitoring instrument, facility assessment and circumcision checklists. Specific items observed included circumcision tasks as accomplished by clinical providers, availability of guidelines in the surgical room, supplies and equipment, infection control, and continuity of care services.

3.4.0 Organization and Implementation of the Study

3.4.1: Sampling and identification of VMMC facilities

Specific considerations informed the sampling process for facilities, procedures and providers as described below.

Facilities providing VMMC: Partner organizations were requested to provide a list of the VMMC sites they supported to carry out VMMCs in Nyanza in 2010 and the number of circumcisions performed from Jan 1 – Dec 31, 2010 at each of the sites. This exercise yielded a list of 235 sites (the study unit of analysis). Out of these, 76 sites were omitted for the following reasons: 12 were sites operated by an organization that would no longer carry out VMMC in Nyanza region in 2011; 29 were "too small" (conducted less than 100 VMMCs in the entire year 2010); 26 were mobile sites that were temporary (i.e., the teams would not visit there a second time but instead move to another location); 9 were reported to be closed or inactive.

This resulted in a final sample frame of 159 sites that met the criterion of performing at least 100 VMMCs in 2010 and representing all four partner organizations conducting VMMC in Nyanza (each partner had a different model of program support for their respective area). The 159 sites were stratified by mode of MMC service delivery model as fixed, outreach and mobile types – described below (mobile sites would only be scheduled by providers where and as service demand dictates). Within each stratum, sites were randomly selected to be representative of the number of VMMCs performed for each service delivery model.

Given a baseline sampling frame of 159 facility units and assuming a response rate of 90% (probability that a unit will be functional on date of visit to enable observation of procedures) among facilities selected (based of previous response rate in other operations research for VMMC in Nyanza) and providing a margin of error of 10%, the sample size required would be as shown in the table below:

Confidence interval-%	90	95	99
Sample size required	22	29	44

Based on the above provisions, considerations of proportionality, within-category distribution, feasibility and resources a stratified random sample size of 30 units (Appendix 1) was deemed sufficient to detect if there was a statistically significant change in circumcision tasks (Bertrand et al., 2014) [15 fixed – category A; and 12 outreach – category B and 3 sites for mobile – category C, which were to be identified once the dates and locations for 2011 operating schedule were known.

In 2012, the intent was to revisit the same 30 sites. However, one of the key implementing partners had drastically reduced the number of its sites, including five of the originally sampled sites. New sites from the original sampling frame that were still operational were determined, and four replacement sites of similar categories from this list were randomly selected. The 5th site was to be a mobile site, identified by a partner organization. However, the Nyanza program had largely shifted from mobile service provision to primarily fixed and outreach service delivery. Thus data collection was not possible at the mobile site. The total number of sites visited was 29, one short of the 30 originally planned.

The sites identified were Ministry of Health approved MC facilities in Nyanza only and were categorized as follows:

1. Specialized or fixed MC sites that offer MC services daily (category A);

2. Outreach MC facilities (category B): These were existing health facilities that incorporate MC into a wider range of services, but do not have an in-house team. Dedicated or part time surgical teams provide services during scheduled clinic day(s) that may be limited to only specific days of the week or hours of the day;
3. Mobile sites (Category C): In these sites surgical teams set up a temporary operating facility in schools or community halls to perform MC on large number of clients following community mobilization.

ii. Sampling of providers: all clinical providers present on the day of visit by researchers and who consented to participate in the study were recruited. It was assumed that at least 90% would agree to participate. This was based on *a priori* experience with other studies on male circumcision using interviews/observation techniques where virtually all providers agreed to participation. We opted for 90% response rate to cater for non-consent by circumcision clients to be observed.

iii. Number of observations: A maximum of 10 observations were made per facility where feasible during the scheduled two day visits. This represents the average number of circumcisions performed per team in a typical facility.

3.4.2 Training: The data collection team consisted of two clinicians (themselves trained in VMMC, to collect data on the clinical aspects of VMMC service delivery) two social scientists. Before commencing data collection, all team members received training from the project PI and Co-PI. The training included: (1) a review of each instrument and the purpose behind each question, (2) general principles of interviewing techniques, (3) specific instructions related to the observation and timing of VMMC procedures, (4) discussion of the flow of activities over the two-day visit at each site, and (5) repeated practice in administering each instrument. Also, the research assistants and their supervisors pre-tested all the data instruments to be used and updates subsequently incorporated. The training also included a substantial component on ethics and informed consent.

3.4.3 Data collection and instruments: Collection of data involved two rounds of actual visits by research assistants to the sites identified in appendix 1. Every site was visited on two consecutive days during each of the rounds in 2011 and in 2012. The visits targeted both the regular and accelerated services in 2011 and 2012 respectively. Only clinical staff available on the day of the visit and consenting to be interviewed and/or observed participated in the study.

Structured questionnaires, observation chart and checklist (Appendices 2, 3 & 4) were used to document information from the sites. (The national male circumcision quality assessment toolkit which is designed to measure progress towards meeting standards and can be used by external assessors to certify or accredit facilities was adapted and modified for the study). Stop-watches were used for counting circumcision time according to protocol.

Quality assessment information was obtained through direct observations of 10 MC procedures per site, where this was feasible and appraising the VMMC facilities using a shortened version of the WHO assessment tool for this purpose. Specifically the measures considered the service setting (29 variables) and circumcision procedure (29 items) (Jennings *et al.*, 2014).

3.5.0 Considerations for Model Construction for DEA and Factor Analysis

3.5.1 Variable selection in constructing DEA model

Although any set of variables may be chosen, the inputs/outputs preferred (defined in table 3.1) closely reflect the organizational context, national program indicators plus the existing functional relationships based on previous DEA-based studies (Kirigia, et al., 1998Hollingsworth, 2008; Jacobs, 2006). Multiple inputs/outputs specification was preferred instead of one to reflect context of delivery context of VMMC intervention (Hacer and Yasar, 2002; Kirigia *et al.*, 2004).

Table 3-1: Table showing input and output variables and their definitions

Index	**Definition of unit of measurement**
Inputs	
Clinician	Number of staff – clinical cadre
Nurse	Number of staff – nursing cadre
Surgical beds	Number of beds in active use
Total operating time (min)	Total elapsed operating time – from entry to exit
Outputs	
MCs performed	Number of circumcisions performed
HTC performed (%)	Proportion of pre-surgical HTC conducted
Quality of service	Service quality index score (based on the average factor scores on a percentile scale)

The criteria as proposed by Bessent and Bessent (1980; 2008) and Hollingsworth was also considered in ascertaining suitability of variable choice:

1. The relationship of inputs to outputs should be realistic.
2. The existing relationship of measured inputs to outputs can be proven.

3. The relationship is such that increases in inputs are associated with increases in outputs.
4. The measurements have no zero elements, but where measurements which have legitimate zero values exist, a small value (0.01) is added to satisfy the model requirement.

The model incorporated quality index variable using methods as suggested by Sherman and Zhu (Sherman & Zhu, 2006) and circumcision time as defined by Rech and colleagues (2014a; 2014b). The quality variable was constructed by factor analysing program data collected during the study. Fifteen process items that correlated highly (conventionally set at ≥0.4) with factor 1 were identified and used to construct quality index for scoring all scheduled tasks per facility. Final index scores were obtained based on the mean score for all tasks considered with higher coefficients representing higher quality on a percentile scale.

Pre-circumcision HIV testing and counselling (HTC) and surgical beds were considered as exogenous variables since providers do not have control over demand for these procedures yet they affect productivity. In particular the number of beds in use is constrained by number of clients and bed-space available at a facility. The number of VMMC providers recorded for each site included only those available on the day of visit by data collection team. The number of circumcisions and percent pre-circumcision HTC performed (as a proportion of all circumcisions at the facility) were estimated by aggregating monthly totals for each site per year starting July, 2010 to July, 2012. All zero observations were substituted with 0.01, since the model does not work with zero (the recommended staff mix for VMMC facilities is to deploy at least two nurses or clinical officers or a combination of both). An output orientation with variable return to scale and 'administrators' perspective' was adopted based on the national VMMC program objective of maximizing outputs, assuming that additional service inputs would result in either proportionate or greater outputs. The input/output variable sets reflect their respective relationship in production.

3.5.2 Sample size for DEA model:

Considerations for sample size were based on:

i) Sample size should be larger than the product of the number of inputs and outputs (Necmi, 2001);

ii) Sample size should be at least 3 times the sum of the number of inputs and outputs (most preferred) (Cooper *et al.*, 2007).

In the current study, the sample comprised of 9 fixed and 13 outreach facilities (9/13). Other study facilities which were visited only once during 2011 or 2012 were excluded from DEA analysis because of the model requirements for comparable number of homogenous facilities at different time points. An additional facility (#106) was dropped from further analysis since no circumcision was observed in 2011. Final total number of facilities analyzed was 21.

3.5.3. Sample size considerations for exploratory factor analysis

With at least 50 variables and factor loadings of at least 0.40 required, the sample size of 58 variables and 369 observations obtained was considered sufficient to produce stable outcomes. Of these 58 items, 4 were omitted due to missing values in some sites. A reasonable ratio for exploratory factor analysis is at least 5:1 (item observations:#variables) (Vyas and Kumaranayake, 2006). The final ratio obtained for this study was 6.8:1.

3.6.0 Statistical Analysis

3.6.1 Data envelopment analysis.

In the first stage analysis, Data envelopment analysis was performed using PIM DEAsoft Ver 3.2, by Ali Emrouznejad and Emmanuel Thennassoulis (2010). Variable relations were determined using Pearson correlations. The service units in the study were evaluated simultaneously to determine and examine their technical efficiency scores and identify those with the most optimal combinations of inputs and outputs (efficient units) relative to their peers in contrast to using an absolute standard. The more efficient ones with regard to their peer unit input utilization vis-à-vis outputs were assigned a value of 100% while those deemed as inefficient were assigned values less than 100% but greater than zero.

In the second stage analysis, sample efficiency estimates were used to test the null hypothesis of no difference in distribution of technical efficiency scores of facilities providing VMMC services between 2011 and 2012 and the decomposed Malmquist productivity indices. Observations for 2011 and 2012 were paired. Paired t-test for two related samples performed to compare means of the paired scores and determine statistical significance of the observed variations under VRS assumption as well as for other model orientations. However, since the statistical properties of data envelopment analysis is still unclear, Wilcoxon sign rank test was also performed to for robustness of the statistical inferences made (Simar and Wilson, 2007).

The following model assumptions were made: The number of circumcisions, surgical beds in use and uptake of pre-operative HTC were considered to be outside the control of providers (exogenous factors) since they depend largely on demand for VMMC. Consequently, the maximum possible increase in outputs was estimated while keeping the inputs and exogenously fixed outputs at their current levels. As demonstrated by Banker and Morey, 1986 (Banker and Morey, 1986) this consideration allows non-discretionary variables to influence the relationship between inputs and outputs, but the "efficiency score" is not affected by them (since they are considered fixed and out of the control of the providers). It also improves comparability of units in the set and enhances opportunities for identifying target increases in the controllable variables required for the facilities to be efficient.

The model orientation and return to scale considered: For the present study technical efficiency was examined under different model versions to elicit the marginal productivity of service units under different assumptions. Specifically, an output orientation with variable returns to scale was considered appropriate since the program aims to maximize outputs within constrained resources and because VMMC facility size (in terms of number of clinical staff and beds used) was deemed relevant to assessing relative efficiency (WHO, 2010a). Returns to scale concept helps to indicate if the facilities are too large or too small considering the inputs used to produce the observed outputs. In the current study bed space, number of staff, uncertain service demand and other exogenous constraints were considered likely to cause VMMC facilities to operate at suboptimal capacities. The efficiency scores obtained from this model version indicate extent of input use for the maximum possible outputs obtained with given unit sizes (Necmi, 2001).

Different model versions make different assumptions about returns to scale: constant returns to scale (CRS), variable returns to scale (VRS) and non-increasing returns to scale (NIRS) (Worthington, 2004). The scale efficiency (SE) is given by the ratio between the CRS and VRS technical efficiency scores (Simar and Zelenyuk, 2006; Zhu, 2003). The VRS assumption ensures that a facility is only compared against others with similar size (based on number of staff and beds) (Necmi, 2001; Zhu, 2003). The VRS model, allows the best practice level of outputs to inputs to vary with the size of the facilities assessed whereas under CRS it is determined by the highest achievable ratio of outputs to inputs for each unit, regardless of size (Kundurjiev, & Salchev, 2011Hollingsworth, 2008). Efficiency scores are identical when computed using input or output orientation under CRS but may vary under VRS. Also, scores obtained when assuming VRS may be higher than or equal to CRS ones since they indicate only technical inefficiency resulting from non-scale factors (Zere *et al.*, 2006).

Pearson correlation was performed on the model variables to examine if they were statistically different. Paired t-test was performed to determine whether there was a statistically significant mean difference in technical efficiency by return to scale: VRS, CRS; NIRS and Scale change in 2011 and in 2012.

Since DEA technical efficiency scores exhibit unknown statistical distribution and that the efficiency scores by CRS, VRS and Scale may be skewed (Cooper, *et al.*, 2007; Spinks and Hollingsworth, 2005), Wilcoxon sign rank test, based on ranking of data was performed to test the robustness of the results.

Malmquist productivity index (MPI): Productivity index measures the change in output-input ratio and indicates shifts in productivity between two time-intervals *(t, t+1)* for each production unit relative to (towards, along or away from) the observed frontier(Hollingsworth, 2008). Malmquist productivity index is one of the methods used to assess productivity changes over time. It identifies the scope and potential sources of production changes based on different types of efficiency measures. Index values >1 implies productivity growth, while a value <1 shows productivity decline, and if =1 indicates stagnation. The estimated index values of production quantities and technological best practice show whether productivity is improving, deteriorating or unchanging over time. These changes are identified by looking into the different types of efficiency change measures: *i)* technical change (associated with variations in quantity and quality of labour / capital); *ii)* pure efficiency change (associated with variations in context / organizational approach largely of labour and capital inputs, including compliance with VMMC treatment protocols and referrals, support supervision, availability of supplies). Both i & ii constitute the overall efficiency change and; *iii)* scale efficiency (attributable to organization changes as unit size varies, for example planning staff mix and work responsibilities, work space and logistics). If there is improved use of resources the service unit position will move towards the frontier indicating positive efficiency gain (Chowdhury *et al.*, 2010).

The Malmquist productivity index was estimated based on Ray and Desli (1997) method in Cooper *et al.*, 2007 (Cooper *et al.*, 2007) to account for scale efficiency change effects as the output mix varied over time (Bert, 2001) with changes in the number of clinical staff and surgical beds used. The average efficiency changes between the two time-periods considered are represented by geometric means. This helps to normalize values because multiple items with different properties are involved.

Statistical tests using Mann Whitney test which is a non-parametric statistical test was performed to examine the mean difference of the total factor productivity, technical efficiency and the respective decomposed components in terms of their positional shift (in relation to the estimated frontier) between 2011 and 2012. It is adopted as proposed by Grosskopf and Valdamanis (1987) and Brockett and Golany (1996) in Cooper et al. (Cooper, *et al.,* 2007) for the non-parametric analysis of DEA results.

Weighting considerations: No *a priori* weight restrictions were imposed on the variables.

Identification of peers: Based on model specifications with exogenous factors fixed, conventional DEA efficiency evaluation of VMMC facilities was performed simultaneously and a reference set of efficient units (peers) identified using a two stage process to ensure identification of both high quality-high efficiency peers. The procedure also identified potential changes required to make each inefficient unit as efficient as the most efficient (best-practice) ones on the frontier (Zhu, 2003). Statistical tests were based on SAS V-13.

3.6.2 Determination of the latent factors of quality assessment toolkit.

Binary interval data from the two instruments for assessing facility preparedness and the other procedure standards, were merged and analyzed using SAS v_13 (SAS Institute Inc. USA). Principal component analysis (PCA) was used for data reduction (and to maximize variability in the data) with factor matrix after standardizing the observed data and developing summary indices, based on the amount of variability accounted for (Floyd and Widaman, 1995; Masaki, 2010; Vyas and Kumaranayake, 2006). In this process, after the first component is defined, consecutive components are extracted from each subsequent residual variance until virtually all variance of measured items are accounted for.

A Scree plot of the eigen values of unrotated factors were plotted to display the graph (Figure 4-2). The steep "cliff" of the curve represents the initial factors extracted and which maximize the variance accounted for, while the shallow “scree” demonstrates small extent of variance accounted for by the subsequent minor factors (Zaslavsky *et al.,* 2002). Conventionally, the cut-off point is where the slope forms an ‘elbow’, being ‘the point at which the slope approaches zero’. Factors with values above this point are retained while those below it are deleted given the variance accounted is almost zero (Floyd and Widaman, 1995; Grimshaw *et al.,* 2003).

Exploratory factor analysis was used to examine the latent structure of extracted components and identify associations among multiple variables comprising each one (Floyd and Widaman, 1995; Landrum *et al.,* 2000).

Rotating factors: to simplify the structure of the variables, Varimax rotation was used since it maximizes the variability of loadings between factors. Simple, meaningful structure is achieved when items load exclusively or highly on as few of the retained factors as possible, but primarily one (Jolliffe, 2002).

Factor loading: The maximum number of iterations was set at 25 to identify variables in each dimension. Factor loadings with absolute values ≥ 0.4 were considered to contribute sufficiently to the overall variability accounted for by the factor (Costello and Osborne, 2005; Stevens, 1992). Cross loading items with values >0.3 were removed to improve consistency.

Quality weights for constructing the index was obtained from the first component as it accounted for the most variability in the items observed (Vyas and Kumaranayake, 2006). The constructed quality index was used to rank the facilities as excellent; good; average and poor by estimating their respective mean scores across the critical quality items. The ranking was based on scores corresponding to the 90th, 75th, 50th and 25th percentiles (0.867; 0.491; -0.219 and -0.667).

3.7 Ethical Considerations

The relevant academic research approval was received from Maseno University School of Graduate Studies. Administrative approval was granted by SYMMACS/USAID to conduct the sub-study. Ethics approval for the SYMMACS study in Kenya was obtained from the Kenya Medical Research Institute (Appendix 5).

Considerations for compliance with the necessary regulatory and administrative requirements in research with human participant were built into the design and implementation of the study. However, there were no substantial risks or harms to the participants at any level during this research. An informed consent (Appendix 6) was administered with every study participant and participating institutions to enhance their autonomy to participate in the interviews or withdraw participation at any time during research. Participants were free to skip questions they did not feel comfortable with and were free to withdraw at any point (two staff declined participation–1 interview; 1-observation). Operational and institutional protocols were duly observed during or prior to site visit.

Audio-/visual-privacy of clients and confidentiality of information received was assured throughout the research process by conducting interviews in secured spaces, keeping the documents securely and coding some of the information, such as identity of sites. Observations of MC procedures were done only after each respective client had consented having been properly briefed by the clinical staff.

CHAPTER 4: RESULTS AND DISCUSSION

4.1.0 Result for Objective 1

4.1.1 Summary Statistics of the Study Variables

DEA Input and output variables: In total there were 21 facilities and 7 variables included in the DEA model (Tables 4-1 and 4-2).

Table 4-1: Facility actual production inputs and outputs in 2011

Year – 2011 Facilities	Input variables				Output variables		
	Clinician	Nurse	Surgical beds	Total elapsed operation time (min)	MCs performed	HTCs % performed	Quality Score
103	1	2	2	50.5	1325	7	1
123	4	3	3	40.05	2488	83.2	60
118	1	1	2	29	1000	68.8	80
102	1	1	2	34.57	137	58	90
125	0	1	1	29.39	414	75	80
126	1	1	4	24.14	256	98	85
108	1	1	3	28.15	785	87.5	40
107	1	1	1	42.2	242	70.8	80
119	1	1	1	28.43	887	84.3	92
104	1	1	2	34.57	137	58	40
111	1	1	4	24.14	256	98	55
110	4	3	3	40.05	2488	83.2	10
112	3	2	4	34.3	1136	86.9	10
129	1	1	1	23.34	176	73	50
124	1	2	1	36.11	516	90	80
101	3	3	1	32.54	2718	66	1
130	1	1	1	39.46	23	80	5
105	1	1	1	23.46	342	74	30
121	0	2	1	46.36	30	0.01	90
114	3	2	4	23.39	2850	99.1	50
109	1	1	1	29.26	688	71	15
Mean	**1.5**	**1.5**	**2.0**	**33.0**	**899.7**	**72.0**	**49.7**
SD	**1.17**	**0.75**	**1.20**	**7.90**	**938.82**	**25.69**	**32.99**

Table legend: Table shows actual data from facilities included in the DEA model. MCs = male circumcisions; HTC = HIV testing and counselling; units used are simple counts. The number of VMMC providers and beds for each site was based on those active/on use on days of data collection visit. Few sites had either zero-clinician or nurse as tasks could be performed by either cadre. Quality scores in percentile.

During 2011, each facility had a mean of 1.5 clinicians and nurses each. Facilities #121 and #125 had nurses only. There was no facility without a nurse. The average quality index score in 2011 was 49.7th percentile. Of all the facilities, 7 (33.3%) scored below 50th percentile (Table 4-1).

In 2012, facilities #105, #118, #121 and #1295 sites had nurses only; while #102, #107, #110 and #111 had clinicians only. The mean quality index score was at 53rd percentile. Of all the facilities 8 (30.1%) scored below 50th percentile (Table 4-2).

Table 4-2: Facility actual production inputs and outputs in 2012

Year 2012 – Facilities	Input variables				Output variables		
	Clinician	Nurse	Surgical beds	Total elapsed operation time (min)	MCs performed	% HTCs performed	Quality Score
103	1	1	1	31.29	1614	86.2	85
123	1	1	1	36.57	1947	85.1	60
118	0	2	2	28.75	837	91.7	30
102	2	0	1	21.38	803	77.1	70
125	2	1	1	28.41	1438	71.7	85
126	1	1	1	35.01	691	83.8	60
108	1	1	1	32.53	376	65	30
107	2	0	1	30.22	1003	62.9	85
119	1	1	1	30.56	764	97.1	25
104	1	1	2	18.17	739	79.8	50
111	2	0	1	16.05	828	66.4	85
110	2	0	1	20.36	63	84.8	50
112	1	2	1	30.28	1430	70.3	85
129	0	2	1	24.19	219	77	15
124	1	2	2	50.5	1325	7	1
101	1	1	1	23.1	2897	97.6	85
130	1	1	1	25.19	1507	99.3	30
105	0	2	1	49.32	552	78.7	15
121	0	2	2	35.42	1516	0.01	40
114	1	1	2	29.13	1616	85	60
109	2	1	1	25.06	243	49.9	70
Mean	**1.1**	**1.1**	**1.2**	**29.6**	**1067.0**	**72.2**	**53.1**
(SD)	**0.70**	**0.70**	**0.44**	**8.75**	**675.59**	**25.96**	**27.40**

Table legend: Table shows actual data from facilities included in the DEA model. MCs = male circumcisions; HTC = HIV testing and counselling; units used are simple counts. The number of VMMC providers for each site was based on those available on days of data collection visit. Few sites had either zero-clinician or nurse as tasks could be performed by either cadre. In the DEA model, the zeroes were substituted with 0.01, since the model does not work with zero. Quality scores in percentile.

Variable slacks: In 2011, no facility had slacks (excess potential) on their clinical officers and nurses input variables, as well as quality index score (an output variable). Facilities # 102, #104, #107 and #111 had both input and output slacks (Table 4-3).

Table 4-3: Facility production Inputs and outputs slacks in 2011

Year - 2011 Facilities	Inputs				Outputs		
	Clinician	Nurse	Surgical beds	Total elapsed operating time (min)	MCs done	HTCs % done	Quality Score
103	0	0	0	0	0	0	0
123	0	0	0	0	0	0	0
118	0	0	0	0	0	0	0
102	0	0	1	6.14	750	26.3	0
125	0	0	0	0	0	0	0
126	0	0	0	0	0	0	0
108	0	0	0	0	0	0	0
107	0	0	0	13.77	645	13.5	0
119	0	0	0	0	0	0	0
104	0	0	1	8.94	122	18.3	0
111	0	0	1	8.94	122	18.3	0
110	0	0	0	0	0	0	0
112	0	0	0	0	0	0	0
129	0	0	0	0	0	0	0
124	0	0	0	0	0	0	0
101	0	0	0	0	0	0	0
130	0	0	0	0	0	0	0
105	0	0	0	0	0	0	0
121	0	0	0	0	0	0	0
114	0	0	0	0	0	0	0
109	0	0	0	0	0	0	0

Table legend: Table shows occurrence of slacks (unused potentials / insufficient outputs) for respective variables. A slack with a positive value indicate need for improvement

In 2012, the following facilities had both input and output slacks: #126, #108, #112, #124, #114 and #109. There was no slack on quality index score variable (Table 4-4).

Table 4-4: Facility production Inputs and outputs slacks in 2012

Year 2012 Facilities	Inputs				Outputs		
	Clinician	Nurse	Surgical beds	Total elapsed operating time (min)	MCs done	% HTCs done	Quality Score
103	0	0	0	0	0	0	0
123	0	0	0	0	0	0	0
118	0	0	0	0	0	0	0
102	0	0	0	0	0	0	0
125	0	0	0	0	0	0	0
126	0	0	0	3.72	923	2.4	0
108	0	0	0	1.24	1238	21.2	0
107	0	0	0	0	0	0	0
119	0	0	0	0	0	0	0
104	0	0	0	0	0	0	0
111	0	0	0	0	0	0	0
110	0	0	0	0	0	0	0
112	0	1	0	7.18	1467	27.3	0
129	0	0	0	0	0	0	0
124	0	1	1	27.4	1572	90.6	0
101	0	0	0	0	0	0	0
130	0	0	0	0	0	0	0
105	0	0	0	0	0	0	0
121	0	0	0	0	0	0	0
114	0	0	1	6.03	1281	12.6	0
109	1	0	0	1.96	2654	47.7	0

Table legend: Table shows occurrence of slacks (unused potentials / insufficient outputs) for respective variables in 2012. A slack with a positive value indicate need for improvement

The estimated correlation coefficients for total elapsed operation time, proportion of HTC and number of circumcisions performed showed statistically significant changes observed across facilities (Table 4-5). Comparison of paired variable mean changes between 2011 and 2012 showed a statistically significant improvement in total elapsed operation time from 33.0 minutes (SD 7.90) to 29.6 minutes (SD 8.75) (95%CI= .0350 – 5.2488; t= 2.114; $df.$= 20; p= 0.047). The mean differences among other variables were not statistically significant.

Table 4-5: Paired Pearson Samples Correlations

Variable pairs: 2011 versus 2012	Correlation	Significance test
Clinical Officer	0.119	0.606
Nurse	0.119	0.606
Surgical beds	0.019	0.934
Proportion of HTC	0.545	**0.011***
Quality score	-0.107	0.644
Total elapsed operation time	0.758	**<0.001***
MCs performed	0.720	**<0.001***

Table legend: Table shows respective level of statistical significance in Pearson correlations between variables as observed in 2011 and 2012. Significant results (Bolded and with star*) indicates the changes observed were consistent among facilities.

4.1.2 Technical Efficiency Analysis by Return to Scale Among Sampled VMMC Facilities

The output orientated technical efficiency scores for all 21 facilities during 2011 and 2012 are shown in table 4-6 below according to dimensions of return to scale.

In 2011 facilities with TE_{VRS}= 100% score were 13 (61.9%) (pure efficiency) and 9 (42.9%) with TE_{CRS}= 100% score (scale efficiency). On the other hand, in 2012 the 16 (76.2%) facilities which were CRS efficient were also VRS efficient (TE_{VRS}= TE_{CRS}= 100%). Five (23.8%) of the remaining facilities were both CRS and VRS inefficient. Thirteen facilities (# 121, 123, 130, 129, 125, 119, 111, 112, 110, 107, 101, 103 and 102) scored 100% during both years while facilities #108, 109, 114, 124 and 126 scored <100% in the same period (Table 4-6).

Table 4-6: Output oriented technical efficiency scores of facilities by year, type, and return to scale (n=21)

Facility type and #		CRS efficiency		VRS efficiency		NIRS Efficiency		Scale efficiency	
		2011	**2012**	**2011**	**2012**	**2011**	**2012**	**2011**	**2012**
Fixed facilities	Fac. 103	75	100	100	100	75	100	75	100
	Fac. 123	57	100	100	100	100	100	57	100
	Fac. 118	82	100	91	100	91	100	90	100
	Fac. 102	100	100	100	100	100	100	100	100
	Fac. 125	100	100	100	100	100	100	100	100
	Fac. 126	92	71	92	71	92	71	100	100
	Fac. 108	43	35	43	35	43	35	100	100
	Fac. 107	97	100	100	100	97	100	97	100
	Fac. 119	100	100	100	100	100	100	100	100
Outreach facilities	Fac. 104	44	100	44	100	44	100	100	100
	Fac. 111	100	100	100	100	100	100	100	100
	Fac. 110	18	100	100	100	18	100	18	100
	Fac. 112	100	100	100	100	100	100	100	100
	Fac. 129	100	100	100	100	100	100	100	100
	Fac. 124	89	1	89	1	89	1	100	81
	Fac. 101	100	100	100	100	100	100	100	100
	Fac. 130	100	100	100	100	100	100	100	100
	Fac. 105	33	100	33	100	33	100	100	100
	Fac. 121	100	100	100	100	100	100	100	100
	Fac. 114	56	71	56	71	56	71	100	100
	Fac. 109	19	82	25	82	19	82	76	100
Summary statistics	**Mean**	**76**	**89**	**84**	**89**	**79**	**89**	**91**	**99**
	SD	**28.7**	**25.2**	**25.3**	**25.1**	**28.9**	**25.1**	**19.8**	**4.0**

Table legend: Table shows the level of resource utilization based on the technical efficiency scores by constant, variable return to scale and Non-increasing Return to Scale; Scale efficiency is a ratio of the CRS : VRS. A difference between the efficiency values under NIRS, VRS and CRS indicate scale inefficiency exists. For example, if $TE_{CRS}<TE_{VRS}$; $TE_{VRS}>TE_{NIRS}$, size is too small, while $TE_{NIRS} = TE_{VRS}$ indicate size too big and at optimal size if $TE_{VRS}=TE_{CRS}$. Additionally, relatively higher standard deviations reflect wider dispersals from the 'best practice' units.

Facilities 102, 107 and 111 achieved 100% technical efficiency but they had slacks in 2011, while facility 112 had slacks in 2012 though technically efficient (table 4-5 and 4-6). The average technical efficiency scores under CRS, VRS and Scale versions were

76% (SD 28.7); 84% (SD 25.3) and 91% (SD 19.8) in 2011 and; in 2012 respective scores were 89% (SD 25.2); 89% (SD 25.1) and 99% (SD 4).

Paired t-test was performed to determine whether there was a statistically significant mean difference in technical efficiency by return to scale: CRS, VRS, NIRS and Scale efficiency during 2011 and 2012. However, the observed differences were not statistically significant, except for scale efficiency (95%CI -31.47959 – 4.698508; $t = -2.8179$; $df. = 20$; $p = 0.005$). Results by the Wilcoxon rank test (W) however indicates that scale efficiency change in 2012 was not different than in 2011 based on the 2-tailed test but the 1-tailed test indicates the difference was significant, as shown in table 4-7.

Table 4-7: Comparison of mean technical efficiency of VMMC facilities (Wilcoxon matched-pairs signed-rank test) n=21

Paired comparisons $H_{O:}$ 2012 > 2011 $H_{A:}$ 2012 < 2011	# observed	Sum of ranks	*z*- value	2-tailed *p* value (Exact)	1-tailed *p* value (Exact)
CRS					
Negative Ranks	3	19.00	-1.569	0.129	0.065
Positive Ranks	9	59.00			
VRS					
Negative Ranks	3	13.00	-.700	0.547	0.273
Positive Ranks	5	23.00			
Scale					
Negative Ranks	1	3.00	-1.859	0.078	**0.039***
Positive Ranks	6	25.00			

Table legend: Table summarizes the statistical inference on distribution of TE scores under different model versions. CRS: constant return to scales; VRS: variable return to scales; *value significant at 99% level. $H_{O:}$ 2012 > 2011; $H_{A:}$ 2012 < 2011; z value= -1.859

Technical efficiency scores by service delivery type: The fixed facilities had a mean efficiency $_{(VRS)}$ of 92% (SD 19.6) in 2011 but declined to 91% (SD 14.5) in 2012 compared to outreach ones with 79% (SD 28.7) in 2011 and 91% (SD 23.0) in 2012.

However the relatively low mean for the outreach facilities could be due to low TE scores in three of the facilities (104, 105, 109) in 2011 and 124 in 2012. The majority of fixed facilities were scale inefficient (TE_{CRS}<TE_{VRS}) in 2011.

The outreach facilities had relatively higher standard deviations reflecting wider dispersals from the mean (Table 4-8). The mean difference of technical efficiency by scale change was statistically significant in favour of outreach facilities (95%CI -45.08035 -2.547979; *df.*= 11; *t*= -2.4647; *p*= 0.015) versus the fixed facilities (95%CI -27.78596 6.874846; *df.*= 8; *t*= -1.3912; *p*= 0.1008). Comparison by variable and constant return to scale showed the difference in means did not reach statistical significance (CRS: 95%CI -9.924384 – 22.92819; *t*= 0.8257; *df.*= 20; *p*= 0.4187; VRS: 95%CI -9.004196 –23.15658; *t*= 0.9179*;* *p*= 0.3696).

Table 4.8: Summary of facility performance scores by type and year

Facility	Statistic	CRS Efficiency		VRS Efficiency		Scale Efficiency	
		2011	2012	2011	2012	2011	2012
Category A - fixed facilities	Mean (%)	83	90	92	90	91	100.0
	SD	21	23	19	23	14.5	0.0
	#100% TE (n= 9)	3	7	6	7	5	9
Category B - Outreach facilities	Mean (%)	77	84	78	84	97	98
	SD	30.7	30.9	29.4	30.9	7.6	5.9
	#100% TE (n=12)	6	9	7	9	10	11

Table legend: Table shows technical efficiency scores of facilities by category under different model versions in 2011 and 2012. SD= standard deviation; TE= technical efficiency; CRS= constant return to scale; VRS= variable return to scale**.**

4.1.3. Discussion for objective 4.1

Of the input variables used, only the elapsed operation time improved significantly by 2012, similar to findings by Rech *et al.,* (2014) in a survey of VMMC program in South Africa. In their study, the reduced operation time did not result in poor service quality, consistent with improved skills. Using DEA it was possible to identify the scope and main sources of overall technical inefficiency change. The observed improvement in the overall mean technical efficiency scores from 84% in 2011 to 89% in 2012 by VRS and 76% to 89% by CRS for the sampled facilities offering male circumcision in Nyanza region shows that facilities still had a scope to improve their output even without altering their inputs by 16% in 2011 and 11% in 2012 under variable return to scales. Inefficiency related to the proportions in which inputs are used or in which outputs are produced in an output orientated DEA model indicates that the technically inefficient facilities largely had excess inputs or insufficient outputs compared to the efficient ones. Observed 'slacks' (unused resources) among technically efficient facilities indicate areas where further improvements are necessary.

Further analysis of the observed distribution of technical efficiency scores under VRS (which expresses only pure technical inefficiency excluding scale factors, and is associated with managerial factors) and CRS, (which expresses global technical inefficiency) during both 2011 and 2012 shows the facilities remained technically inefficient. The significant improvement in scale efficiency scores during 2012 is an indication of improved relative unit size, in terms of the ratio of input-output mix (the number of input resources used by facilities vis-à-vis outputs produced). While unit sizes generally improved, managerial challenges related to issues such as compliance to standard operating procedures persisted in 2012.

Previous DEA evaluation of health care delivery at various delivery tiers in Kenya by Kirigia *et al.,* (2004; 1998) showed inefficiencies related to resource use and insufficient outputs. However, no empirical studies on technical efficiency of specific vertical intervention programs implementation in Kenya *per se* have been conducted before.

The practical application of these results for MC managers is that improvement efforts should focus on managerial aspects and capacities that augment the specific outputs (service quality, number of circumcisions and proportions of circumcised men tested for HIV) and minimize inputs used, including the number of staff and the elapsed operation time. Resource use and service output at fixed facilities can be addressed through policy and operational guidelines as well as skills augmentation.

4.2.0 Results for Objective 2

4.2.1 Identifying Peers and Benchmark Units for the Observed VMMC Facilities in Nyanza Based on the DEA Model

The facilities with efficiency score of 100% exhibit ideal performance. Those identified by DEA to be peers or "role-models" for the less-than efficiently performing facilities appear in the efficiency reference set (ERS). Table 4.9a shows an output orientated VRS results of inefficient facilities and their peers with respective combination weights in parenthesis.

Table 4-9a: Initial DEA results showing inefficient units, their corresponding efficiency reference sets and relative weight respectively assigned to each

Year/DMU		Efficiency score	Reference facility (relative weights)	Peer count
2011	108	43%	Fac. 119 (1)	3
	126	92%	Fac. 119 (1)	
	109		Fac. 129 (0.78) Fac. 121 (0.22)	2
	105	33%	Fac. 121(1)	4
	114	56%	Fac. 121(1)	
	118	91%	Fac. 125(0.32) Fac. 119 (0.68)	1
	104	44%	Fac. 121 (1)	-
	124	89%	Fac. 111 (1); Fac. 129 (0.84)	1
2012	126	71%	Fac. 103 (1)	2
	108	35%	Fac. 103 (1)	
	112	112%	Fac. 101 (1)	4
	124	1%	Fac. 101 (1)	
	114	71%	Fac. 101 (1)	
	109	82%	Fac. 101 (1)	

Table legend: Table shows initial relative peer assessment of facility efficiency scores. DMU= Decision-making unit; fac.= facility; Peer count/freq= number of times the facility is a peer; appropriate combination weights (λ) required to enable respective facilities reach relative efficiency are in parenthesis.

These values show projected production options that will enable them reach relative efficiency. Facilities identified as peers were #111, 119, 121, 129 and 125 in 2011; 103 and 101 in 2012. However, the reference facilities #129 and 111 had high technical efficiency scores but low in quality score (50 and 55 respectively). The DEA model was reran but excluding the 2 low-quality units to enhance probability of obtaining only high efficiency-high quality peers, following recommendations by Sherman and Zhu (2006) and, Shimshak, Lenard and Klimberg (2009).

The resulting new reference units shown in table 4-9b were obtained. All had higher efficiency scores and quality values; facility 103 had a quality score of 60 but high performance in other output variables. Facilities 119 and 125 dominated in 2011 and 103 in 2012.

Table 4-9b: Revised DEA model results after deleting low quality facilities: inefficient units, their corresponding efficiency reference sets and relative weight respectively assigned to each

Year/DMU		Efficiency score	Reference facility(relative weights)	Peer count
2011	118	91%	Fac. 119 (0.68); Fac. 125 (0.32)	- 1
	108	43%	Fac. 119 (1);	3
	126	92%	Fac. 119 (1)	
2012	126	71%	Fac. 103 (1);	2
	108	35%	Fac. 103 (1)	

Table legend: Table shows repeat peer assessment of facilities with low quality ones (#129 and #111) excluded. DMU= Decision-making unit; fac.= facility; peer count/frequency count is the number of times the facility appear as a peer; figures in parenthesis are appropriate combination weights required to enable respective facilities attain efficiency

4.2.3 Discussion for Objective 4.2

The study has demonstrated that consideration of service quality in identifying peers for benchmarking VMMC facilities helped to improve discrimination of ideal benchmark units. This finding is similar to previous observation (Park *et al.,* 2012) which shows that using multiple criteria including preference structure, direction and similarity in choosing DEA variables improves its precision in benchmarking and hence potential for ensuring balance and comprehensiveness of evaluation decisions.

However, when considering service quality variable for inclusion in DEA benchmarking Shimshak *et al.,* (2009) recommend that “the choice of quality output measures be appropriately related to the input measures” to improve compatibility with the objectives of the DEA model. Furthermore, Sherman and Zhu (2006) have observed an efficiency/quality trade-off when benchmarking with quality-adjusted DEA to seek lower-cost-high quality service in the banking industry.

In the current study, the only three 'best practice' peers identified in the efficiency reference set were from fixed facility category. This could be attributed to the fact that outreach service types had unique and diverse characteristics in terms of size and operational dynamics compared to the fixed types. This implies that when planning improvement efforts, it is necessary for managers to consider the contextual needs for each facility and other occult causes of inefficiencies unique to them despite their position in relation the frontier. This can be accomplished by stratifying the facilities by their preference structure, operational direction and similarity. These parameters define the progress of a facility towards attaining its efficiency target (Park *et al.,* 2012).

Since no precedence exist in performing productivity analysis or regulatory benchmarking of health facilities providing VMMC care in Kenya, these findings demonstrate the potential and desirability of incorporating service quality into benchmarking of health service delivery units to enhance optimal decisions.

4.3.0 Results for Objective 3

4.3.1 Characterize Changes in Productivity in VMMC Service Delivery Function Based on DEA Malmquist Productivity Index

Productivity performance change by Malmquist Index: Table 4-10 shows there was progress in observed total factor productivity (83.4%), mainly attributable to improvement in technical change (72.9%).

Table 4.10: Productivity performance for each service delivery facility by type

Facility code	TC	SEC	PEC	TFPG (MI)	Efficiency-2011	Efficiency 2012
Fac. 103	1	1.16	1	1.16	100	100
Fac. 123	1	1.33	1	1.33	100	100
Fac. 118	0.96	1.18	1.1	1.24	90.75	100
Fac. 102	1	1	1	1	100	100
Fac. 125	1	1	1	1	100	100
Fac. 126	0.93	1	0.76	0.71	92.39	70.59
Fac. 108	0.63	1	0.81	0.51	43.48	35.29
Fac. 107	1	1.02	1	1.02	100	100
Fac. 119	1	1	1	1	100	100
Fac. 104	0.97	1	2.25	2.19	44.44	100
Fac. 111	1.14	1	1	1.14	100	100
Fac. 110	2.92	2.34	1	6.83	100	100
Fac. 112	2.92	1	1	2.92	100	100
Fac. 129	1.3	1	1	1.3	100	100
Fac. 124	8.96	0.22	0.01	0.03	88.89	1.18
Fac. 101	1	1	1	1	100	100
Fac. 130	4.12	1	1	4.12	100	100
Fac. 105	0.97	1	3	2.92	33.33	100
Fac. 121	1	1	1	1	100	100
Fac. 114	1.16	1	1.27	1.47	55.56	70.59
Fac. 109	1.32	1.08	3.23	4.62	25.49	82.35
Mean	**1.729**	**1.063**	**1.211**	**1.834**	**84.492**	**88.571**
SD	**1.822**	**0.345**	**0.716**	**1.586**	**25.326**	**25.120**

Table Legend: TC: technical change (frontier shift); SEC: scale efficiency change; PEC: pure efficiency change. PEC + SEC = efficiency change (catching up effect); TFPG (MI): total factor productivity growth (Malmquist Index). The table shows individual productivity values vary widely. Total factor productivity growth improved by 83.4%. The key driver of technical efficiency was technical change (72.%) attributable to new service delivery technologies and labor.

Pure efficiency change improved by 21.1% and scale efficiency change by 6.3%. (Table 4.10). Thirteen (61.9%) of the VMMC facilities were efficient during the period. Of these, facilities #103, 123, 118, 104, 107, 110, 111, 112, 124, 130, 105, 114, 109 & 129 experienced total productivity growth and three others #108, 124 & 126 (19.0 %) regressed (Table 4-10 above).

Assessment results of the observed changes for statistical significance using the non-parametric Mann Whitney U test for independent data indicate that the observed progress for total factor productivity by 83.4% and technical change by 72.9% were statistically significant. However, there was insufficient evidence that the observed distribution for the other components of technical efficiency were significantly different, as shown in table 4-11.

Table 4-11: Mann Whitney U test for the Malmquist productivity index (n=21)

Variable	Facility category	Sum of Ranks	Mann-Whitney U	Wilcoxon W	Z	*p-Value* (2-tailed)	*p-Value* (2-tailed)
Technical change (TC)	Fixed facilities Outreach facilities	63.00 168.00	18.000	63.000	-2.633	**0.008**	0.009
Scale efficiency Change (SEC)	Fixed facilities Outreach facilities	115.50 115.50	37.500	115.500	-1.397	0.162	0.247
Pure efficiency change (PEC)	Fixed facilities Outreach facilities	82.00 149.00	37.000	82.000	-1.383	0.167	0.247
Total Productivity Factor Growth (TPFG)	Fixed facilities Outreach facilities	69.00 162.00	24.000	69.000	-2.147	**0.032***	0.034

Table legend: there were 9 fixed facilities; 12 0utreach facilities. Table shows the observed technical efficiency change distribution across facilities were significant for TPFG and TC.

Fixed facilities experienced significant decline in total factor productivity growth (TPFG) change and technological change (TC) (0.3% / 5.3%) but decline in pure efficiency change (PEC) by 3.7% and progress in scale efficiency by 2.3% were not statistically significant. On the other hand, the outreach ones progressed in all factors of productivity although wider dispersals from the mean were observed among them (Table 4-12).

Table 4-12: Productivity indices for VMMC facilities by types between 2011 and 2012

Facility Category / Statistic		**TC**	**SEC**	**PEC**	**TFPG**	**2011 (TE %)**	**2012 (TE %)**
Category 'A' (fixed)	Mean	0.947	1.077	0.963	0.997	91.847	89.542
	SD	0.114	0.113	0.101	0.240	17.447	21.260
Category 'B' (outreach)	Mean	2.315	1.053	1.397	2.462	78.976	87.843
	SD	2.235	0.445	0.899	1.854	28.690	27.641

Table legend: Table shows productivity indices of facilities by category. TC: Technological change (Boundary shift due to technological change); SEC: Scale Efficiency Change; PEC: Pure technical Efficiency Change; TFPG: Total Factor Productivity Growth (Malmquist Index); Values equal to 1 implies no productivity change; <1 decline and, >1 progress in productivity. TE: technical efficiency score (%)

4.3.2 Discussion for Objective 4.3

The main driver of productivity increase, largely among outreach facilities, was technological change associated with improved ‘speed’/experience in performing circumcisions and a reduction in the number of input resources deployed. This is similar to a study by Rech *et al.,* (2014a) who observed a significant relationship between mean operating time and quality score (in south African program set up) indicating improvement in provider experience and skills. A descriptive comparison of facilities across four countries by Jennings et al. (2014) using the same data-set showed a statistically significant decline in the number of quality tasks performed correctly. This could explain lack of statistical significance in the mean quality score index in the current study.

On the other hand, the observed scale efficiency changes, only contributed slightly to productivity growth indicating that improved facility size was not a major factor in productivity growth. Also, lack of statistically significant change in pure technical efficiency implies the operational context between the two years was not sufficiently different, which may imply that managerial circumstances did not significantly vary over the period. This is plausible considering the widespread poor compliance to operational standards among clinical providers as well as effects of exogenous factors.

The decline in technical efficiency and factor productivity among fixed VMMC facilities was attributable mainly to technological and pure inefficiency which reflect probable adverse challenges related to their operating environments, staff skills and other institutional management factors. These facilities are unlikely to adjust optimally to demand changes quickly due to inelasticity in obligatory resources, especially labour and theatre-space (Afonso and Fernandes, 2008). It is not clear however how this may change over a long period of time when appropriate methods are adopted.

Comparatively, the outreach service delivery is a more viable delivery method for enhancing VMMC program coverage. However, as saturation point for coverage is achieved, sub-optimal resource use may set in. Both methods of service delivery were beset with organizational challenges. Consequently the Ministry of Health policies and implementing organization managements could seek to emphasize improvements of operational contexts through service integration with investment in staff management skills in addition to augmenting technical skills. This is more critical for fixed facility service delivery methods when considering mainstreaming VMMC for long-term sustainability while de-emphasizing roles of outreach service delivery strategy (Hirschberg and Lye, 2001; Wang, 2009).

4.4.0 Result for Objective 4

4.4.1 Latent Factor Dimensions and Variable Interrelationships of the System-Based VMMC Service Quality Assurance Tool Using Factor Analysis Techniques

4.4.1.1 Principal component analysis.

Principal Component Analysis (PCA) as a data reduction technique was used to extract key components from the observed data, based on their eigen values. A total of 54 item measures and 246 responses with normal distribution were analyzed. Based on a stepwise approach such that the first few factors successively account for most of the variation in the original observed set of variables, 58 common factors (latent variable) for the measures were identified. The initial estimate of common variance among all the 58 factors was 45, accounting for 77.6% of the total variance. Fifteen factors with eigen values ≥ 1 accounted for 73.6% of the total variance. Based on Eigen value >1.00 (Floyd and Widaman, 1995) three factors, each respectively with Eigen values of 5.78; 4.29; 2.99 were retained (Figure 4-1). These factors cumulatively accounted for 29.1% of the total variance (12.9%; 9.5%; 6.7%) with final communality estimates being 13.06.

Figure 4-2: Scree plot showing distribution of factors by their eigen values

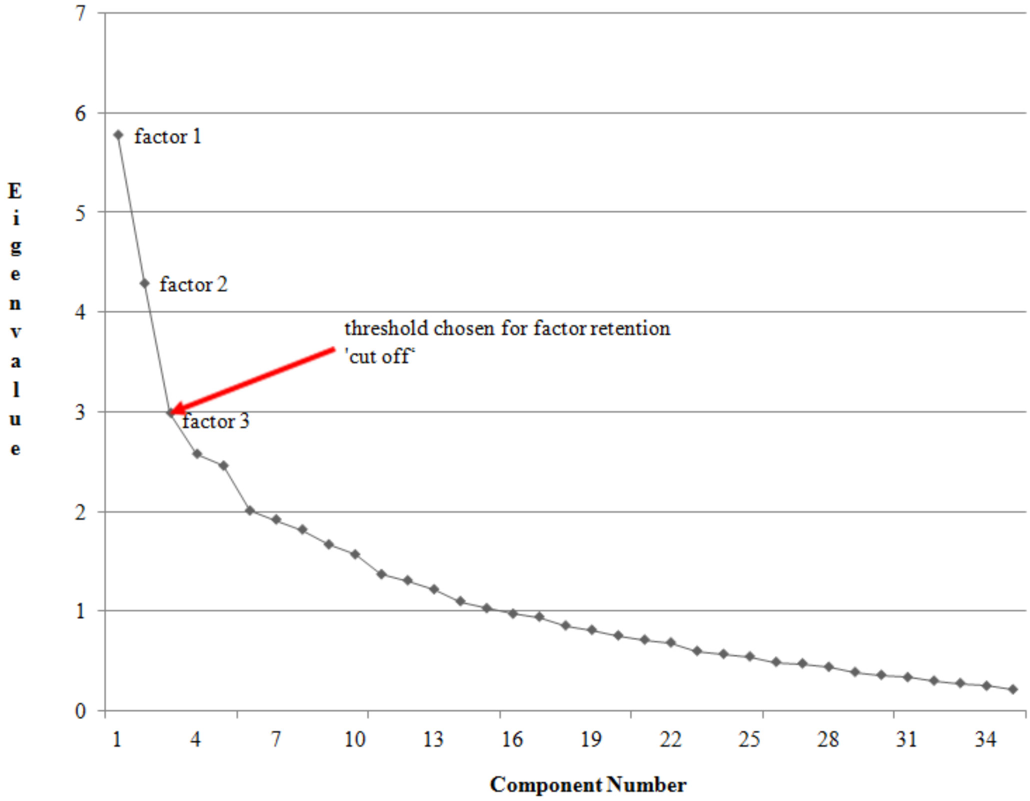

Figure 1: Scree plot showing distribution of factors by their eigenvalues

Figure 4-1 legend: A Scree plot of eigenvalues of the unrotated factors displaying an 'elbow' of the plot (shown by the red arrow). This point of the curve represents the threshold chosen for retention of the initial factors extracted from the observed variables and which maximize the variance accounted for. Three factors, each respectively with eigenvalues of 5.78; 4.29; 2.99 were retained. These factors cumulatively accounted for 29.1% of the total variance (12.9%; 9.5%; 6.7%) with final communality estimates being 13.06. The shallow "scree" distal to the arrow demonstrates small extent of variance accounted for by the subsequent minor factors, which were deleted.

4.4.1.2 Exploratory factor analysis.

Respective factor items with eigen values ≥0.4 are displayed list-wise in Table 4-13. Based on this cut off, fifteen (15) items loading on factor 1, five (5) on factor 2 and one (1) on factor 3 were retained. Factor one items relate to different aspects of VMMC service delivery quality indicators, focusing broadly on the preparedness to provide VMMC in terms of physical infrastructure, guidelines and the interactive elements of circumcision service. Hence it is labelled 'preparedness to deliver safe male circumcisions'. Items converging on factor 1 can be categorized further into: *safety reliability* (availability of basic life support equipment, eligibility assessment, observation of vital signs and other events post-operatively to identify potential harms and, availability of antibiotics for treatment of adverse events); *appropriateness* (using guidelines in performing necessary pre-, intra- and post-operative tasks); *communication interaction* (pre- and post- operative information-giving on HIV and circumcision); *access to minimum service package* (syndromic management of STIs, individualized confidential HTC and condom distribution) and *staff competence* (correct surgical knots tying technique).

Factor 2 is labelled 'performance-safety' being related to skill-compliance issues and safety of surgical procedure. It comprises of variables related to *continuity of care* (discharge care and interactive follow-up instructions); *staff safety* (eye wear to prevent splash to the eyes); *acute care* (oxygen as a basic life support). The only item loading on factor 3 retained based on the threshold for cut off was *'Appropriate antibiotics in stock to treat infection related AEs'*. However, it was also cross-loading on factor 1. Since Factor three had only one item, it was considered weak hence was not included as part of the analyses.

Table 4-13: Rotated factor loadings of factor 1, 2 and 3 relating to VMMC service quality dimensions

Variables	Factor 1 [preparedness]	Factor 2 [performance -safety]	Factor 3
Basic life support equipment (CPR) is on hand in case of an emergency -Antihistamine.	**0.67**	0.16	0.30
Staff reviews vital signs.	**0.64**	0.32	-0.04
Staff observes post-op clients for an allergic reaction or any other abnormality before allowing them leave the operating table or recovery room.	**0.62**	0.33	-0.37
Basic life support equipment (CPR) is on hand in case of an emergency:- IV Lines.	**0.60**	0.13	0.39
Appropriate antibiotics in stock to treat infection related AEs.	**0.58**	-0.15	**0.44**
Basic life support equipment (CPR) is on hand in case of an emergency:-Bag &mask.	**0.57**	**0.51**	0.30
WHO guidelines for performing MC or National guideline of Standard Operating Procedure (SOP) for VMMC available on site.	**0.53**	-0.35	0.34
Clinical personnel conduct a basic preoperative examination.	**0.52**	-0.13	0.09
Patients receive post -operative counseling instructions and reinforcement of previous MC/HIV messaging.	**0.51**	0.36	-0.23
National protocols for syndromic management of STIs available on site.	**0.51**	-0.20	0.38
Facility offers HIV testing and counseling (HTC).	**0.49**	-0.73	-0.22
Staff provides private individual counseling and question time on VMMC and offers HTC.	**0.49**	-0.73	-0.22
Male condoms available for distribution to clients.	**0.48**	-0.15	-0.27
Staff provide patients with clear instructions, (verbal and written) on how to wash and care for the wound and how to deal with pain and minor bleeding	**0.46**	0.25	-0.21
Uses correct technique in tying surgical knots.	**0.41**	0.21	-0.35
Basic life support equipment (CPR) is on hand in case of an emergency: - Oxygen supply.	-0.23	**0.71**	0.31
Use of protective eyewear by all providers.	0.29	**0.51**	0.30
Staff gives specific reminders of the 6 week post of the 6 week post-operative abstinence period.	-0.11	**0.46**	-0.25
Staff insist/encourage clients to return for at least one follow up visit or in case of a complication.	0.29	**0.51**	-0.47

Table legend: Three principal components retained after varimax rotation of the PCA factors. Only factor coefficients with values ≥0.4 and minimal cross-loading were considered valuable. Factor 3 was finally omitted from further consideration because of significant cross-loading with factor 1 of the only item loading on it with large value.

4.4.2 Discussion for Objective 4.4

The current study was undertaken from the management perspective to enable objective evaluation of provider performance. This contrasts with most existing studies on health service quality which largely focus only on the consumers' viewpoint, and are less likely to reflect accurately aspects like provider competence (Babikako *et al.,* 2011; Parasuraman, Zeithaml, & Berry 1985; Tabrizi, 2011; Zaneta and Ilona, 2008). The factor analysis of the SYMMACS quality instrument reveals three main discreet factors. The value of extracted factors in measuring service quality is however contingent on the observed data and the relationships between variables under consideration as well as the validity and reliability of the variables retained (Jolliffe, 2002). Likewise, rotation of factors to simplify structure may cause loss of variance on individual dominant sources (Jolliffe, 2002).

The factor loadings show the hierarchical item importance within the factors, in terms of both component availability and task performance. The observed variability in quality of VMMC service delivery is best explained by the dimensions *'preparedness to deliver safe male circumcisions'*, being complemented by *'performance-safety'*. Implicit in these dimensions are the technical and functional requirements necessary for accomplishing VMMC service delivery tasks correctly.

The SYMMACS quality instruments exhibit similar underlying concepts as those described in existing health service quality studies (Casey *et al.,* 2009; Donabedian, 1988; Hong *et al.,* 2006; Shahidzadeh-Mahani *et al.,* 2008; Sower *et al.,* 2001). The principal factors also configure well with the WHO toolkit criteria (WHO, 2009) for assessing VMMC service quality and the domains (Effective, Appropriate, Safe, Efficient, Responsive, Accessible, Continuous, Capable, Sustainable) in other health systems quality performance frameworks (Landgren and Murray, 2008; World Health Organization, 2006). The principal component elicited represents diversity of service quality sources (Landrum *et al.,* 2000) and reflects the multidimensionality of VMMC services similar to other public health interventions.

The factor structure elicited specifically demonstrates safety aspects and provider-client interactions as key quality considerations. This would potentially assist program managers to understand the scope encompassed by VMMC quality assessment tool, recognize its importance and progressively build into the delivery system capacities for proper service performance. The observations also indicate that staff performed dismally mainly in the tasks related to spontaneous patient-staff communication interactions, particularly engagement in the post-operative period in contrast to availability of equipment and supplies which by default are provided by the program (Jennings *et al.*, 2014). These performance tasks are inherently related to individual competency and responsiveness which if emphasized would greatly improve patient safety. This is consistent with findings by Malmquist index above.

As part of quality improvement plans, VMMC operational guidelines ought to subsequently clarify systemic approaches to health practice safety. Likewise, emphasizing compliance with operational guidelines will ensure desired service quality outcomes are obtained (WHO, 2009; Zaneta and Ilona, 2008). Refresher staff training and support supervision are helpful in enhancing progressive learning of skills for target tasks and responsibilities besides monitoring how well these are performed (van Duong *et al.,* 2004; Zaneta and Ilona, 2008). These capacities includes communication abilities and interpersonal skills to improve information-giving (Sofaer and Firminger, 2005).

A service quality model developed by Parasuraman *et al.* (1985) has 10 dimensions: tangibles, reliability, responsiveness, competence, courtesy, credibility, assurance, access, communication, and customer understanding. Brown and Swartz (1989) applied this model to assess quality of medical-surgical service delivery and they determined that this list is reasonably applicable to health service settings. Bruce-Jain's framework however, consists of six dimensions: needs assessment, choice of contraceptive methods, information given to users, interpersonal relations, constellation of services, and continuity mechanisms. This has been used to assess contraceptive services to adolescents in Uganda (Nalwadda *et al.*, 2011). Given the dimensions apparent in the SYMMACS service quality instruments are both reasonable and actionable, a simplified version based on the principal factors can be adopted for routine quality assessment and monitoring.

4.5.0 Results for Objective 5

4.5.1 Ranking Service Quality Performance of Circumcisions in Nyanza Region Using Quality Index Comprising the Critical Items of the Quality Toolkit

A composite index measure comprising the critical items of the quality toolkit was constructed from the fifteen variables with coefficient values between 0.4 – 0.7 after varimax rotation of factor 1, since it contained most of the variations observed in the original data set. These weighted factor coefficients were used to rank facilities and all the cases that were performed (Table 4-13).

Table 4-14: Overall facility ranking by weighted quality scores

Facility Identification #	Average facility index value	Percentile
111	1.325	≥ 90th percentile [1= excellent]
101	1.288	
112	1.246	
133	1.244	
103	1.174	
107	1.168	
125	1.098	
134	1.059	
131	0.980	
102	0.429	75th percentile [2= Good]
109	0.417	
123	0.275	
114	0.223	
126	0.201	
104	0.155	
106	-0.053	
110	-0.065	
136	-0.107	
121	-0.432	50th percentile [3= Average]
130	-0.563	
118	-0.564	
132	-0.645	
108	-0.774	≤ 25th percentile [4= poor]
119	-0.980	
137	-1.103	
129	-1.201	
105	-1.500	
124	-2.230	

Table legend: the composite index tool is based on and reflects the respective item loadings on factor 1, which contained most of the variations observed in the original data set (conventional minimum> 60%). Facility values (column 2) were calculated as the average of service quality scores across all the items in the assessment list. Ranking on a percentile scale ranged from excellent (1) to poor (4) (column 3)

Respective facility values were calculated as the average of service quality scores across all the critical items in the assessment list, converted to a percentile scale.

Ranking on a percentile scale ranged from excellent (1) to poor (4). Out of the 28 facilities included for factor analysis, 32% (9/28) had scores between 90th and 95th percentile while 45% had percentiles between 50th and 75th. In four of the facilities, scores for all cases observed were poor, being in the lower 25th percentile.

Based on the composite quality scores, 50% of the VMMC cases that performed in 2011were ranked as good or excellent compared to 58.8% in 2012, while almost a quarter of them were ranked as poor (Table 4-14).

Table 4-15: Ranking of cases using composite service quality index by year of study

Ranking	Number of cases by year			
	2011		2012	
	Freq.	%	Freq.	%
Excellent	22	24.4	71	32.6
Good	23	25.6	57	26.2
Average	22	24.4	41	18.8
Poor	23	25.6	49	22.5
Total	**90**	**100**	**218**	**100**

Table legend: Ranking was based on the average value of the composite index score for each case observed per facility.

4.5.2 Discussion on Objective 4.5

Overall more than half of the facilities and cases assessed scored showed scale-up has been largely satisfactory. However the occurrence of low service quality scores in some of the facilities raises serious concerns about their capacity to comply with recommended guidelines as program scale up continues. In South Africa, Rech *et al.* (2014b) found similar country level trends in VMMC service quality. They report challenges with both the technical and functional dimensions of quality as the program expanded. These observation call for more concerted efforts to enhance acquaintance with practice criteria among service providers as well as interventions to improve support supervision. While it is not known how long these inherent challenges in scaling up a complex program as VMMC would last as the implementation of rapid expansion continues, to mitigate adverse effects on service quality in the long-term, it is desirable to avail sufficient resources and guidelines to facilitate facility preparations as well as for correctly accomplishing scheduled tasks. In addition conducting support supervision consistently to enhance compliance with established standards is essential.

4.6 Limitations of the study

Due to inability to revisit some of the facilities in the second year of study, as a result of program shifts, a number of facilities were dropped from DEA analysis due to DEA requirements for comparable units. This reduced the sample size, and it is not known the impact of excluding the observations. The inclusion of composite quality output variable still requires further research on how best it should be included in DEA techniques. Also, since it was not possible that all model specifications be considered the interpretations reached may not be generalizable to other models.

Furthermore, since the statistical characteristics of DEA outputs are still hazy, the statistical significance results and inferences made on technical efficiency and productivity changes must be treated with caution. With regard to quality assessment, lack of the client perspective is a limitation of this study, although it is still unknown how its inclusion would alter the characteristics of the derived factor constructs, given that theoretically, this aspect as an outcome is difficult to link with the structure and process that produce it unless comprehensive reliable information is available.

CHAPTER 5: SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

5.1 Summary of Findings

The study results have shown that the DEA-based average technical efficiency of VMMC service delivery was 97.1% in 2011 and 88.9% in 2012. However, there was no sufficient evidence to show that this change was statistically significant. On the other hand, there was a significant change in scale efficiency, indicating improvement in facility size (by bed-space and number of staff). The facilities dominating as peers provide appropriate benchmarks for modelling service operations and improvement objectives for other facilities in the reference set to work towards. Based on the Malmquist factor productivity index it was shown that technological progress was responsible for improved productivity growth. However, the observed inefficiency was mainly attributable to managerial challenges given the decline in pure technical efficiency change, this was not statistically significant. Consequently, program improvement efforts should focus on building resource management capacities and institutional reorganization to deal with underutilized resources and enhancing technical skills. A composite index constructed from items clustering around the latent principal component of the quality assessment toolkit was used to score performance of cases in terms of technical and functional dimensions. Only slightly more than half of the circumcision procedures performed ranked above average. Based on the twenty item measures of the two latent components, appropriate model scoring may be constructed and adopted for routine program assessment of the dimensions which are required for accomplishing VMMC service delivery tasks correctly.

5.2 Conclusions

The conclusions of the study are as follows:

1. The efficiency scores by return to scale showed there was improvement in overall mean scale technical efficiency consistent with the observed reduction in inputs and increased outputs. However, pure technical inefficiency persisted over the period, indicating managerial/operational challenges.

2. Using DEA techniques the study showed that benchmarking is a useful strategy to objectively identify technically efficient facilities (using optimal resource combinations. This is helpful in setting performance targets ideal for the inefficient service delivery facilities.

3. Based on DEA Malmquist productivity index, the sources of improved factor productivity variations in VMMC service delivery in this sample was due to technological efficiency, which is mainly associated with technical skills. This is consistent with the observed improved total elapsed operation time.

4. Using exploratory factor analysis techniques, it was determined that the latent factor structure of the service quality monitoring toolkit comprise of three dimensions which broadly explains its construct. Of the items in the toolkit, only 20 explain the majority of observed variations of the assessment outcome. They are closely related to the technical and functional VMMC service tasks. They are the critical quality measure items in VMMC service delivery.

5. Cases ranked based on the composite index derived from factor analysis indicated that the proportion performing above average was just slightly above half.

5.3 Recommendations from the Study

The following are the recommendations from the study based on the results:

1. The technical efficiency scores by return to scale showed persisting pure technical inefficiency (associated with managerial factors). It is recommended that VMMC service providers should be trained on both technical and managerial aspects of care delivery to enhance optimal resource/space allocation. In addition, there is need to maintain optimal input resource size relative to the output.

2. Evaluating and benchmarking facility performance is an objective way of identifying best performing facilities within a complex multidimensional health intervention context. The DEA techniques provide valuable tool for program managers to use in this case. It is recommended that peer benchmarking be used routinely to set improvement targets for the less efficient/poor performing facilities service delivery units.

3. Based on the factors of productivity in VMMC service delivery, outreach service delivery approach had higher factor productivity gain, hence may be preferred while fixed-facility approach remains complementary. With regard to programme efficiency improvement needs, structural and organizational adjustments to improve pure technical efficiency and scale technical efficiency are necessary in some of the facilities experiencing space and managerial constrains. Additional training is required to enable managers within implementing institutions to improve on resource management skills. The national program strategic documents should clarify specific requirements for

each service function parameter to ensure reasonable balance in both technical efficiency and scale/coverage efficiency.

4. Since the factor structure of the service quality monitoring toolkit was clarified and simplified in this study, it is recommended that the national program adopts a simpler quality measure instrument comprising of only 20 critical quality items out of the 54 items currently used for quality assessment.

5. A composite quality index aggregating all these variables would potentially be simpler for routine use for ranking service quality performance of VMMC facilities. In addition, field supervision is required to support compliance with operational guidelines on communication, performance of safe procedures and interpersonal skills to ensure desired service outcomes are obtained.

5.4 Recommendations for Future Studies

Recommendations for further studies include:

1. Service delivery function: Qualitative studies of service delivery function are necessary to determine specific exogenous factors that are more likely to adversely influence technical efficiency
2. DEA performance evaluation: Studies are needed to explore time lag effects on VMMC service function activities where the period from use of inputs to production of outputs occurs in the long-term post-surgical period.
3. Composite quality index: Further studies should explore how best composite quality index measures can be included into the DEA model as variable.

4. Further quality studies are necessary to explore development of appropriate composite service quality measures for VMMC service delivery using item dimensions based on larger routine data sets.

REFERENCES

Adams, O., Shengelia, B., Stilwell, B., Larizgoitia, I., Issakov, A., Kwankam, S.Y., and Jam, F.S.T. (2003). Provision of personal and non-personal health services: Proposal for monitoring. Murray CJL, Evans DB. Health systems performance assessment: Debates, Methods and Empiricism. Ginebra: Organizacion Mundial de la Salud *WHO/EIP/OSD/DP 02.25*, 235-249.

Afonso, A., and Fernandes, S. (2008). Assessing hospital efficiency: Non-parametric evidence for Portugal. Available at SSRN 1092135.

Akazili, J., Adjuik, M., Chatio, S., Kanyomse, D., Hodgson, A., Aikins, M., and Gyapong, J. (2008). What are the technical and allocative efficiencies of public health centres in Ghana? Ghana Med J *42*, 149 - 155.

Alexander, C.A., Busch, G., and Stringer, K. (2003). Implementing and interpreting a data envelopment analysis model to assess the efficiency of health systems in developing countries. IMA Journal of Management Mathematics *14*, 49-63.

Arah, O.A., Klazinga, N.S., Delnoij, D.M.J., Ten Asbroek, A.H.A., and Custers, T. (2003). Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement. International Journal for Quality in Health Care *15*, 377-398.

Auvert, B., Taljaard, D., Lagarde, E., Sobngwi-Tambekou, J.l., Sitta, R.m., and Puren, A. (2005). Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial. PLoS Med *2*, e298, 1112 - 1122.

Babikako, H.M., Neuhauser, D., Katamba, A., and Mupere, E. (2011). Patient satisfaction, feasibility and reliability of satisfaction questionnaire among patients with pulmonary tuberculosis in urban Uganda: a cross-sectional study. Health Research Policy and Systems *9*, 6.

Bailey, R.C., Egesah, O., and Rosenberg, S. (2008). Male circumcision for HIV prevention: a prospective study of complications in clinical and traditional settings in Bungoma, Kenya. Bulletin of the World Health Organization *86*, 657-736.

Bailey, R.C., Moses, S., Parker, C.B., Agot, K., Maclean, I., Krieger, J.N., Williams, C.F., Campbell, R.T., and Ndinya-Achola, J.O. (2007). Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. Lancet *369*, 643-656.

Banker, R.D. (1993). Maximum likelihood, consistency and data envelopment analysis: a statistical foundation. Management science *39*, 1265-1273.

Banker, R.D., and Morey, R.C. (1986). The use of categorical variable in data envelopment analysis. Management Science *32*, 1613-1627.

Bellows, J., and Sullivan, M.P. (2004). Background Paper: Could a Quality Index Help Us Navigate The Chasm? In “Facilitating Use of Clinical Quality Measures through Aggregation” (Washington, DC. , Kaiser Permanente; Institute for health policy).

Bert, M.B. (2001). Scale Efficiency and Productivity Change. Journal of Productivity Analysis *15*, 159-183.

Bertrand, J.T., Rech, D., Aduda, D.O., Frade, S., Loolpapit, M., Machaku, M.D., Oyango, M., Mavhu, W., Spyrelis, A., and Perry, L. (2014). Systematic monitoring of voluntary medical male circumcision scale-up: adoption of efficiency elements in Kenya, South Africa, Tanzania, and Zimbabwe. PloS one *9*, e82518. DOI: 82510.81371/journal.pone.0082518.

Bessent, A.M., and Bessent, E.W. (1980). Determining the comparative efficiency of schools through data envelopment analysis. Educational Administration Quarterly *16*, 57-75.

Bollinger, L., DeCormier, P., and Stover, J. (2009). Male Circumcision: Decision Makers’ Program Planning Tool, Calculating the Costs and Impacts of a Male Circumcision Program. (Washington, DC, Futures Group, Health Policy Initiative, Task Order 1).

Breen, A.M., Burton-Houle, T., and Aron, D.C. (2002). Applying the theory of constraints in health care: Part 1--The philosophy. Qual Manag Health Care *10*, 40-46.

Burgess, J.F. (2012a). Innovation and efficiency in health care: does anyone really know what they mean&quest. Health Systems *1*, 7-12.

Burgess, J.F. (2012b). Innovation and efficiency in health care: does anyone really know what they mean[quest]. In Health Systems (Operational Research Society Ltd), p. 12.

Campbell, S.M., Roland, M.O., and Buetow, S.A. (2000). Defning quality of care. Social Science & Medicine *51*, 1611 -1625.

Casey, S., Mitchell, K., Amisi, I., Haliza, M., Aveledi, B., Kalenga, P., and Austin, J. (2009). Use of facility assessment data to improve reproductive health service delivery in the Democratic Republic of the Congo. Conflict and Health *3*, 12.

Chen, L., Evans, T., Anand, S., Boufford, J.I., Brown, H., Chowdhury, M., Cueto, M., Dare, L., Dussault, G., and Elzinga, G. (2004). Human resources for health: overcoming the crisis. The Lancet *364*, 1984-1990.

Chisholm, D., and Evans, D.B. (2010). Improving health system efficiency as a means of moving towards universal coverage. In World health report, 2010 (Geneva, World Health Organization), pp. (1 - 33).

Chowdhury, H., Zelenyuk, V., Wodchis, W., and Laporte, A. (2010). Efficiency and Technological Change in Health Care Services in Ontario. (Centre for Efficiency and Productivity Analysis).

Cooper, W.W., Seinford, M.L., and Tone, K. (2007). Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA Solver Software, Second Edition edn (New York: Kluwer Academic Publishers).

Costello, A.B., and Osborne, J.W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. Practical Assessment Research & Evaluation *10*.

Dan, C., and David, B.E. (2010). Improving health system efficiency as a means of moving towards universal coverage. (Geneva, Switzerland, World Health Organization).

Derose, S.F., and Petitti, D.B. (2003). Measuring quality of care and performance from a population health care perspective. Annual Review of Public Health *24*, 363-384.

Derose, S.F., Schuster, M.A., Fielding, J.E., and Asch, S.M. (2002). Public health quality measurement: Concepts and Challenges. Annu. Rev. Public Health *23*, 1-21.

Donabedian, A. (1988). The quality of care - how can it be assessed. Jama-Journal of the American Medical Association *260*, 1743-1748.

Eklund, F.J. (2008). Resource constraints in health care-case studies on technical, allocative and economic efficiency. In Industrial Engineering and Management (Espoo, Finland, Helsinki University of Technology), p. 143.

Färe, R., Grifell-Tatjé, E., Grosskopf, S., and Lovell, C.A.K. (1997). Biased Technical Change and the Malmquist Productivity Index. Scandinavian Journal of Economics *99*, 119-128.

Fernando, S.S., and Abeynayak, R. (2012). Modified Factor Analysis to Construct Composite Indices: Illustration on Urbanization Index Tropical Agricultural Research *23*, 327-337.

Floyd, F.J., and Widaman, K.F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment *7*, 286-299.

Gray, R.H., Kiwanuka, N., Quinn, T.C., Sewankambo, N.K., Serwadda, D., Mangen, F.W., Lutalo, T., Nalugoda, F., Kelly, R., Meehan, M., *et al.* (2000). Male circumcision and HIV acquisition and transmission: cohort studies in Rakai, Uganda. Rakai Project Team. AIDS *14*, 2371-2381.

Gray, R.H., Li, X., Kigozi, G., Serwadda, D., Nalugoda, F., Watya, S., Reynolds, S.J., and Wawer, M. (2007). The impact of male circumcision on HIV incidence and cost per infection prevented: a stochastic simulation model from Rakai, Uganda. AIDS *21*, 845-850.

Grimshaw, J., McAuley, L.M., Bero, L.A., Grilli, R., Oxman, A.D., Ramsay, C., Vale, L., and Zwarenstein, M. (2003). Systematic reviews of the effectiveness of quality improvement strategies and programmes. Qual Saf Health Care *12*, 298 - 303.

Hacer, O., and Yasar, A.O. (2002). A National Study of Efficiency for Dialysis Centers: An Examination of Market Competition and Facility Characteristics for Production of Multiple Dialysis Outputs. Health Services Research *37*, 711 - 732.

Hankins, C., Forsythe, S., and Njeuhmeli, E. (2011). Voluntary medical male circumcision: an introduction to the cost, impact, and challenges of accelerated scaling up. PLoS medicine *8*, e1001127. DOI: 1001110.1001371/journal.pmed.1001127.

Hathorn, E., Land, L., and Ross, J.D.C. (2011). How to assess quality in your sexual health service. Sexually Transmitted Infections *87*, 508-510.

Herman-Roloff, A., Bailey, R.C., and Agot, K. (2012). Factors associated with the safety of voluntary medical male circumcision in Nyanza Province, Kenya. Bull World Health Organ *90*, 773 - 781.

Herman-Roloff, A., Llewellyn, E., Obiero, W., Agot, K., Ndinya-Achola, J., Muraguri, N., and Bailey, R.C. (2011). Implementing Voluntary Medical Male Circumcision for HIV Prevention in Nyanza Province, Kenya: Lessons Learned during the First Year. PloS one *6*, e18299. DOI: 18210.11371/journal.pone.0018299.

Hirschberg, J.G., and Lye, J.N. (2001). Clustering in a data envelopment analysis using bootstrapped efficiency scores. (Department of Economics, University of Melbourne), pp. 1 - 22.

Hollingsworth, B. (2003). Non-parametric and parametric applications measuring efficiency in health care. Health Care Manag Sci *6*, 203-218.

Hollingsworth, B. (2004). Non Parametric Efficiency Measurement. The Economic Journal. *114*, F307-F311.

Hollingsworth, B. (2006). Measuring Efficiency in Health Care: Analytic Techniques and Health Policy. By ROWENA JACOBS, PETER C. SMITH and ANDREW STREET. Economica *77*, 205-206.

Hollingsworth, B. (2008). The measurement of efficiency and productivity of health care delivery. Health economics *17*, 1107-1128.

Hollingsworth, B., Dawson, P.J., and Maniadakis, N. (1999). Efficiency measurement of health care: a review of non-parametric methods and applications. Health Care Manag Sci *2*, 161-172.

Hong, R., Montana, L., and Mishra, V. (2006). Family planning services quality as a determinant of use of IUD in Egypt. BMC Health services research *6*, 79. doi:10.1186/1472-6963-1186-1179.

Hulten, C.R. (2001). New Developments in Productivity Analysis. In Total Factor Productivity. A Short Biography, Charles R. Hulten, E.R. Dean, and M.J. Harper, eds. (University of Chicago Press), pp. p. 1 - 54.

Hulten, C.R., Dean, E., and Harper, M. (2001). Front matter, New Developments in Productivity Analysis. In New Developments in Productivity Analysis, Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, eds. (Chicago: University of Chicago Press), pp. 17 - 20.

Issel, L.M. (2004). Health Program Planning and Evaluation: A Practical, Systematic Approach for Community Health Jones & Bartlett Publishing. London.

Jacobs, R., Smith P.C., Street A. (2006). Measuring Efficiency in Health Care - Analytic Techniques and Health Policy.

Jennings, L., Bertrand, J., Rech, D., Harvey, S.A., Hatzold, K., Samkange, C.A., Aduda, D.S.O., Fimbo, B., Cherutich, P., and Perry, L. (2014). Quality of voluntary medical male circumcision services during scale-up: a comparative process evaluation in Kenya, South Africa, Tanzania and Zimbabwe. PloS one *9*, e79524. DOI: 79510.71371/journal.pone.0079524.

Jolliffe, I.T. (2002). Principal Component Analysis, Second Edition. In Springer Series in Statistics, P.D. P. Bickel, S. Fienberg, K. Krickeberg, I. Olkin, N. Wermuth, S. Zeger, ed. (London, Springer).

Kahn, J.G., Marseille, E., and Auvert, B. (2006). Cost-effectiveness of male circumcision for HIV prevention in a South African setting. PLoS Med *3*, e517.

Kirigia, J.M., Emrouznejad, A., and Sambo, L.G. (2002). Measurement of Technical Efficiency of Public Hospitals in Kenya: Using Data Envelopment Analysis. Journal of Medical Systems *26*, 39 - 45.

Kirigia, J.M., Emrouznejad, A., Sambo, L.G., Munguti, N., and Liambila, W. (2004). Using Data Envelopment Analysis to Measure the Technical Efficiency of Public Health Centers in Kenya. Journal of Medical Systems *28*, 155 - 166.

Kirigia, J.M., Fox-Rushby, J., and Mills, A. (1998). A cost analysis of Kilifi and Malindi public hospitals in Kenya. Afr J Health Sci *5*, 79-84.

KNBS, I.-M. (2010). Kenya Demographic and Health Survey 2008-09. (Calverton, Maryland, Kenya National Bureau of Statistics (KNBS) and ICF Macro).

Kounetas, K., and Tsekouras, K. (2007). Measuring Scale efficiency change Using a translog distance function. International School of Business and economics *6*, 63 - 69.

Kredo, T., Gerritsen, A., van Heerden, J., Conway, S., and Siegfried, N. (2012). Clinical practice guidelines within the Southern African development community: a descriptive study of the quality of guideline development and concordance with best evidence for five priority diseases. In Health Research Policy and Systems.

Kundurjiev, T., and Salchev, P. (2011). Technical effiency of hospital psychiatric care in Bulgaria: assessment using Data Envelopment Analysis. Online at http://mpra.ub.uni-muenchen.de/28953/. In Munich Personal RePEc Archive (MPRA) (Department of Social Medicine and health care management).

Landgren, F., and Murray, J. (2008). A guide to using data for health care quality improvement. Victorian Quality Council Secretariat, ed. (Big Print, 50 Lonsdale St, Melbourne VIC 3000, Rural and Regional Health and Aged Care Services Division, Victorian Government Department of Human Services, Melbourne, Victoria).

Landrum, M.B., Bronskill, S.E., and Normand, S.T. (2000). Analytic methods for constructing cross-sectional profiles of health care providers. Health Serv Outcomes Res Methodol *1*, 23 - 47.

Lazar, E.J., Fleischut, P., and Regan, B.K. (2013). Quality Measurement in Healthcare. Annual Review of Medicine *64*, 485-496.

Leah, F.B., and Barbara, R. (2009). Commentary: A Systematic Review of Health Care Efficiency Measures. Health Services Research *44*, 806-811.

Lee, S.T., and Holland, D. (2001). The impact of Noisy Catch data on estimates of efficient output derived from DEA and Stochastic Frontier Models: a Monte Carlo Comparison. In International Institute of Fisheries Economics and Trade (IIFET) (Corvallis, Oregon, U.S.A), pp. 1 -9.

Lukas, C.V., and Hall, C. (2010). Challenges in Measuring Implementation Success. In 3rd Annual NIH Conference on the Science of Implementation and Dissemination: Methods and Measurement.

Mahler, H.R., Kileo, B., Curran, K., Plotkin, M., Adamu, T., Hellar, A., Koshuma, S., Nyabenda, S., Machaku, M., and Lukobo-Durrell, M. (2011). Voluntary medical male circumcision: matching demand and supply with quality and efficiency in a high-volume campaign in Iringa Region, Tanzania. PLoS medicine *8*, e1001131. DOI: 1001110.1001371/journal.pmed.1001131.

Mangham, L.J., and Hanson, K. Scaling up in international health: what are the key issues? Health Policy and Planning.

Mangione-Smith, R., DeCristofaro, A.H., Setodji, C.M., Keesey, J., Klein, D.J., Adams, J.L., Schuster, M.A., and McGlynn, E.A. (2007). The Quality of Ambulatory Care Delivered to Children in the United States. New England Journal of Medicine *357*, 1515-1523.

Masaki, M. (2010). How to factor-Analyze your data right: Dos and Don'ts. International Journal of Psychological Research *3*, 97 - 110.

McGlynn, E.A., Shekelle, P.G., Chen, S., Goldman, D.P., Romley, J.A., Hussey, P.S., de Vries, H., Wang, M.C., Timmer, M.J., and Carter, J. (2008). Identifying, categorizing, and evaluating health care efficiency measures. Final Report (prepared by the Southern California Evidence-based Practice Center--RAND Corporation, under Contract No. 282-00-0005-21). AHRQ Publication No. 08-0030. (Rockville, MD, Agency for Healthcare Research and Quality), p. 124.

MoH (2005). Reversing the Trends: The Second National Health Sector Strategic Plan (NHSSP II, 2005 - 2010). (Nairobi).

MOPHS (2011). Ministry of Public Health and Sanitation. Provincial MMC Task force report, Nyanza Province; April, 2011.

Murray, C.J.L., and Frenk, J. (1999). A WHO framework for health system performance assessment (Geneva: Evidence and Information for Policy, World Health Organization).

Mwandi, Z., Murphy, A., Reed, J., Chesang, K., Njeuhmeli, E., Agot, K., Llewellyn, E., Kirui, C., Serrem, K., and Abuya, I. (2011). Voluntary medical male circumcision: translating research into the rapid expansion of services in Kenya, 2008â€“2011. PLoS medicine *8*, e1001130. DOI: 1001110.1001371/journal.pmed.1001130.

NACC (2009). Kenya National AIDS Strategic Plan 2009/10 - 2012/13: Delivering on Universal Access to Services. (Nairobi, Ministry of Public Health and Sanitation).

Nagelkerke, N.J.D., Moses, S., de Vlas, S.J., and Bailey, R.C. (2007). Modelling the public health impact of male circumcision for HIV prevention in high prevalence areas in Africa. BMC infectious diseases *7*, 16. doi:10.1186/1471-2334-1187-1116.

Nalwadda, G., Tumwesigye, N.M., Faxelid, E., Byamugisha, J., and Mirembe, F. (2011). Quality of Care in Contraceptive Services Provided to Young People in Two Ugandan Districts: A Simulated Client Study. PloS one *6*, e27908. doi: 27910.21371/journal.pone.0027908.

NASCOP (2008a). Manual for Male Circumcision under Local Anaesthesia, Version 2.5C. (Nairobi, National AIDS and STI Control Programme).

NASCOP (2008b). National Guidance for Voluntary Male Circumcision in Kenya. (Nairobi, Ministry of Health, Republic of Kenya).

NASCOP (2009). Kenya National Strategy for Voluntary Medical Male Circumcision. (Nairobi, Ministry of Public Health and Sanitation).

National AIDS & STI Control Program (2009). The National Policy on Voluntary Medical Male Circumcision in Kenya. (Nairobi, Kenya).

National AIDS/STI Control Programme (October 2010). Voluntary Medical Male Circumcision for HIV Prevention in Kenya: Report of the First Rapid Results Initiative, November/December 2009. (Nairobi, Kenya, Government of Kenya, Ministry of Public Health and Sanitation,).

Naveen, D., and Boonghee, Y. (1998). Retail Productivity Assessment Using Data Envelopment Analysis. Journal of Retailing *74*, 89-105.

Necmi, K.A. (2001). Investigating technical and scale efficiencies of Australian Universities through data envelopment analysis. Socio-Economic Planning Sciences *35*, 57 - 80.

Nietert, P., Wessell, A., Jenkins, R., Feifer, C., Nemeth, L., and Ornstein, S. (2007). Using a summary measure for multiple quality indicators in primary care: the Summary QUality InDex (SQUID). Implementation Science *2*, 11.

Njeuhmeli, E., Forsythe, S., Reed, J., Opuni, M., Bollinger, L., Heard, N., Castor, D., Stover, J., Farley, T., Menon, V., and Hankins, C. (2011). Voluntary Medical Male Circumcision: Modeling the Impact and Cost of Expanding Male Circumcision for HIV Prevention in Eastern and Southern Africa. PLoS Med 8(11) *8*, e1001132. doi:10.1371/journal.pmed.1001132

Obiero, W., Young, M.R., and Bailey, R.C. (2011). The PrePex device is unlikely to achieve cost-savings compared to the forceps-guided method in male circumcision programs in sub-Saharan Africa. PloS one *8*, e53380. DOI: 53310.51371/journal.pone.0053380.

Parasuraman, A., Zeithaml, V.A., and Berry, L.L. (1985). A conceptual model of service quality and its implications for future research. Journal of marketting *49*, 41 - 50.

Park, J., Bae, H., and Lim, S. (2012). A DEA-Based Method of Stepwise Benchmark Target Selection with Preference, Direction and Similarity Criteria. International Journal of Innovative Computing, Information and Control *8*, 5821 - 5834.

Rapkin, B., Weiss, E., Chhabra, R., Ryniker, L., Patel, S., Carness, J., Adsuar, R., Kahalas, W., DeLaMarter, C., Feldman, I., *et al.* (2008). Beyond satisfaction: Using the Dynamics of Care assessment to better understand patients' experiences in care. Health and Quality of Life Outcomes *6*, 20.

Rech, D., Bertrand, J.T., Thomas, N., Farrell, M., Reed, J., Frade, S., Samkange, C., Obiero, W., Agot, K., Mahler, H., *et al.* (2014a). Surgical Efficiencies and Quality in the Performance of Voluntary Medical Male Circumcision (VMMC) Procedures in Kenya, South Africa, Tanzania, and Zimbabwe. PloS one *9*, e84271. DOI: 84210.81371/journal.pone.0084271.

Rech, D., Spyrelis, A., Frade, S., Perry, L., Farrell, M., Fertziger, R., Toledo, C., Castor, D., Njeuhmeli, E., Loykissoonlal, D., and Bertrand, J.T. (2014b). Implications of the Fast-Evolving Scale-Up of Adult Voluntary Medical Male Circumcision for Quality of Services in South Africa. PloS one *9*, e80577. DOI: 80510.81371/journal.pone.0080577.

Reinhardt, U.E., and Tsung-mei, C. (2000). The world health report 2000-Health systems: improving performance. Bulletin of the World Health Organization *78*, 1064-1064.

Shahidzadeh-Mahani, A., Omidvari, S., Baradaran, H., and Azin, S. (2008). Factors affecting quality of care in family planning clinics: A study from Iran. International Journal for Quality in Health Care *20*, 284-290.

Shaller, D. (2004). Implementing and Using Quality Measures for Children’s Health Care: Perspectives on the State of the Practice. Pediatrics *113*, 217-227.

Sherman, H.D., and Zhu, J. (2006). Benchmarking with quality-adjusted dea (q-dea) to seek lower-cost high-quality service: Evidence from a U.S. bank application. . Annals of Operations Research *145*, 301-319.

Shimshak, D.G., Lenard, L.M., and Klimberg, R.K. (2009). Incorporating Quality into Data Envelopment Analysis of Nursing Home Performance: A Case Study. Omega. *37*, 672-685.

Simar, L., and Wilson, P.W. (2007). Estimation and inference in two-stage, semiparametric models of production processes. J Econ *136*, 31 - 64.

Simar, L., and Zelenyuk, V. (2006). On testing equality of distributions of technical efficiency scores. Econometric Reviews *25*, 497-522.

Smith, P.C., and Street, A. (2005). Measuring the efficiency of public services: the limits of analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society) *168*, 401-417.

Sofaer, S., and Firminger, K. (2005). Patient perceptions of the quality of health services. Annual Reviews in Public Health *26*, 513-559.

Sower, V., Duffy, J., Kilbourne, W., Kohers, G., and Jones, P. (2001). The dimensions of service quality for hospitals: development and use of the KQCAH. Heathe Care Management Review *26*, 47 - 59.

Spinks, J., and Hollingsworth, B. (2005). Health production and the socioeconomic determinants of health in OECD countries: the use of efficiency models (Monash University. Centre for Health Economics).

Steering Committee for the Review of Commonwealth / State Service Provision (1997). Data envelopment analysis: A technique for measuring the efficiency of government service delivery. (Commonwealth of Australia, AGPS, Canberra).

Stevens, J.P. (1992). Applied Multivariate Statistics for the Social Sciences (2nd edition) (Hillsdale, NJ: Erlbaum.).

Tabrizi, J. (2011). Quality of Delivered Care for People with Type 2 Diabetes: A New Patient-Centred Model. Journal of Research in Health Sciences, North America *9*, 1 - 9.

Torkian, E., and Ranjbar, H. (2013). Health technical efficiency estimation of sub-Saharan africa. The international Journal of Social Science *12*, 11.

UNAIDS (2009). AIDS Epidemic Update. (Geneva, Joint United Nations Programme on HIV/AIDS).

Union, A. (2005). Africa Health Strategy: 2007 - 2015. (Abuja, Nigeria).

van Duong, D., Binns, C.W., Lee, A.H., and Hipgrave, D.B. (2004). Measuring client-perceived quality of maternity services in rural Vietnam. International Journal for Quality in Health Care *16*, 447-452.

Veihola, M. (2008). Technical efficiency of blood component preparation in blood centres of 10 European countries. In Faculty of Medicine, Department of Public Health, Finnish Red Cross Blood Service (Helsinki, University of Helsinki), p. 64.

Vyas, S., and Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. Health Policy and Planning *21*, 459-468.

Wegbreit, J., Bertozzi, S., DeMaria, L.M., and Padian, N.S. (2006). Effectiveness of HIV prevention strategies in resource-poor countries: tailoring the intervention to the context. AIDS *20*, 1217-1235.

Weiss, H.A. (2007). Male circumcision as a preventive measure against HIV and other sexually transmitted diseases. Curr Opin Infect Dis *20*, 66 - 72.

WHO (2000). The World Health Report 2000: Health Systems: Improving Performance.

WHO (2004). Strengthenin health systems:The role and promise of policy and systems research (Geneva: Alliance for health policy and systems research).

WHO (2009). Male Circumcision Services: Quality Assessment Toolkit. (WHO Geneva, Switzerland).

WHO (2010a). Considerations for implementing models for optimizing the volume and efficiency of male circumcision services. Field testing edition.

WHO (2010b). The relative efficiency of public and private service delivery. In World Health Report: Background Paper, 39: The path to universal coverage: Health Systems Financing, J. Hsu, ed. (Geneva).

WHO/UNAIDS (2007). WHO and UNAIDS announce recommendations from expert consultation on male circumcision for HIV prevention. (Media center).

WHO/UNAIDS (2008). Male Circumcision and HIV Prevention in Eastern and Southern Africa: Communications Guidance.

WHO/UNAIDS (2009). A guide to indicators for male circumcision programmes in the formal health care system. World Health Organization and Joint United Nations Programme on HIV/AIDS 2009, ed. (Geneva, Switzerland, WHO Document Production Services).

World Health Organization (2006). Quality of care: a process for making strategic choices in health systems. p. 38 p.

World Health Organization (2007). Everybody's business--strengthening health systems to improve health outcomes: WHO's framework for action. (Geneva, World Health Organization), p. 44.

Worthington, A.C. (2004). Frontier Efficiency Measurement in Health Care: A Review of Empirical Techniques and Selected Applications. Medical Care Research and Review *61*, 135-170.

Yanshuang, Z., and Byungho, J. (2012). A DEA Model for Performance Evaluation in The Presence of Time Lag Effect. World Academy of Science, Engineering and Technology *69*, 517 - 520.

Ying-Ming, W., Ying, L., and Liang, L. (2009). Ranking decision making units by imposing a minimum weight restriction in the data envelopment analysisI. Journal of Computational and Applied Mathematics *223*, 469-484.

Zaneta, P., and Ilona, B. (2008). Different perspectives on health care quality: Is the consensus possible? Engeneering economics *1*, 104 - 110.

Zaslavsky, A.M., Shaul, J.A., Zaborski, L.B., Cioffi, M.J., and Cleary, P.D. (2002). Combining health plan performance indicators into simpler composite measures. Health Care Financ Rev *23*, 101 - 115.

Zere, E., Mbeeli, T., Shangula, K., Mandlhate, C., Mutirua, K., Tjivambi, B., and Kapenambili, W. (2006). Technical efficiency of district hospitals: evidence from Namibia using data envelopment analysis. Cost Eff Resour Alloc *4*, 5.

Zhu, J. (2003). Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets International Series in Operations Research Management Science. (Worcester Polytechnic Institute, Department of Management, Worcester, MA 01609).

APPENDICES

Appendix 1: Study Sites – 2011 & 2012

Site	Type	District	Implementing Partner
Bondo D.Hosp	A	Bondo	NRHS
Mawere	B	Siaya	NRHS
Usigu H/Center	B	Bondo	NRHS
Okok	B	Ndhiwa	APHIA Plus
UNIM	A	Kisumu East	NRHS
Chulaimbo RHTC	A	Kisumu West	NRHS
Manywanda	A	Kisumu West	NRHS
Macalder D. H	A	Nyatike	FACES
Winjo Disp	B	Nyatike	IRDO
Lwanda disp	B	Nyatike	IRDO
Ndhiwa D.Hosp	A	Ndhiwa	IRDO
TYC Ahero	A	Nyando	IRDO
St Vincent's Mission Hospital	B	Nyando	IRDO
Kadinda Center	B	Kisumu East	IRDO
Kandiege SDH	A	Homa Bay	APHIA Plus
Adiedo Disp	A	Rachwonyo	NRHS
Misambi Disp	A	Rachwonyo	NRHS
Othoro Health Center	A	Rachwonyo	APHIA Plus
Madiany SDH	A	Rarieda	NRHS
Awendo SDH	A	Rongo	APHIA Plus
Ramula dispensary	B	Siaya	NRHS
Malanga dispensary	B	Siaya	NRHS
Tingare Health Center	B	Siaya	NRHS

Mbita District Hospital	A	Mbita	FACES
TYC Suba	A	Mbita	IRDO
Ogongo Health Center	A	Suba	FACES
Nyatoto Health Center	B	Suba	IRDO
Rapiedha Community center	C	Ndhiwa	IRDO
Nyamaiya H/C	C	Nyamira	NRHS
St Francis Asisi	C	Rachwonyo	NRHS
Chiga Community center	C	Kisumu East	IRDO
Arcon Hosp	B	Ndhiwa	APHIA Plus
Magunga Health Center	A	Suba	IRDO
TYC Central Kisumu East	A	Kisumu East	IRDO
Koru Mision Hospital	B	Nyando	IRDO
Sony Sugar Medical Center	A	Rongo	APHIA Plus
Lagnet Community Resource & Information Center	C	Nyando	IRDO

Appendix 2: SYMMACS INSTRUMENT #1-A: Characteristics of the Male Circumcision Facility

Name of site: ________________________________
Code for site: ______ Date: ____DD __MM __YY
Name/code of clinical observer: _____________________
Name or code of site manager providing data: ____

Instruction: the clinician obtains the data for the following chart from the site manager or other person responsible for the site on the day of the visit:

	Day 1	**Day 2**
Number of beds in use on day of visit		

Number by cadre of primary provider(s) performing MC		
Physician		
Assistant Medical Officer (AMO)		
Clinical officer		
Nurse		
Number by cadre of secondary provider(s) assisting with MC:		
Physician		
Clinical officer		
AMO (where applicable)		
Nurse		
Number of non-medical assistants that clean and organize surgical area (hygienist, runner, cleaner, etc.)		
Total number of MC procedures performed		

The remainder of instrument #1 is based on (1) interviewing the chief medical administrator at the MC facility, and (2) confirming through visual inspection the presence of data, supplies, and equipment on site.

#	**Items to be observed and scored:**	**0 = none**	**1= partial**	**2= total**	**Comments**
CHARACTERISTICS OF THE FACILITY:					
	Adequate lighting in surgical area				
	Adequate ventilation in surgical area				
	General appearance of MC facility (including surgical area) – clean, hygienic				
AVAILABILITY OF DATA (manual or computerized files)					
	Existence of a functioning				

	information system that collects: date of operation, client's name, age, procedure performed, anesthesia given, surgeon's name, comments				
	Consent forms on file for every client circumcised on the day of visit.				
	Monitoring system in place for adverse effects (on the day or at follow-up) that records: patient's name, ID #, nature and severity of adverse effect, and treatment of AE				
#	**Items to be observed and scored:**	**0 = none**	**1= partial**	**2= total**	**Comments**
AVAILABILITY OF ESSENTIAL ITEMS ON SITE					
	WHO guidelines for performing MC or National guideline of Standard Operating Procedures (SOP) for MC				
	Sterilized instruments available for use during MC				
	Local anesthesia (correctly stored, not expired)				
	Antibiotics in stock to treat infection related AEs				
	Pain medication in stock				
	Antiseptic solution in stock				
	Dressing materials (bandages and gauze) in stock				
	Basic life support equipment (CPR) is on hand in case of an emergency: -- Bag and mask for CPR				
	--Oxygen supply				

	--IV lines and resuscitation fluids				
	--Antihistamine, cortisone and adrenalin to treat anaphylaxis				
	HIV post exposure prophylaxis in stock				
	Guidelines available on site for post exposure prophylaxis in stock				
	Sharps container available in surgical area				
	National protocols for syndromic management and treatment of STIs available				
	Male condoms available for distribution to clients				
	Facility offers HIV counseling and testing (HCT)				
#	**Items to be observed and scored:**	**0 = none**	**1= partial**	**2= total**	**Comments**
	Facility area that provides visual and auditory privacy for HCT and disclosure of results				

	PREOPERATIVE PROCEDURES:				
	Staff provides group education on risks and benefits of MC surgery including behavior change counseling.				
	Staff provides private individual counseling and question time on MC and offers HCT.				

	Site has referral slips for clients requiring other services (including those with contraindications for MC)				
	SUPERVISORY MECHANISM IN PLACE				
	Site manager reports receiving a supervisory visit in past 6 months.				
	Site manager reports that an external source has monitored the Adverse Effects rate of the program within the past 6 months.				

Comments (optional):

___

___

(READ ALOUD): I'd like to discuss a few issues related to demand creation and client load.

1. At this site do you have (***READ THE RESPONSES***):
   ___too many clients for the operating capacity
   ___too few clients (you could do more MCs per day if you had more clients)
   ___a good balance between number of clients and your ability to provide MC
   INSTRUCTION TO INTERVIEWER: DON'T READ "IT DEPENDS," BUT IF THE SITE MANAGERS GIVES THIS ANSWER, THEN TICK "IT DEPENDS" AND ASK HIM/HER TO EXPLAIN:
   ___ it depends (***EXPLAIN***):
   ___
   ___

2. My final question relates to demand creation activities to encourage male circumcision in this catchment area (that is, the population served by this site). These may be activities organized at the national level (such as radio or TV) or activities organized by your own site (such as mobilization). To the best of your knowledge, which of the following communication channels have been used to promote MC in your catchment area within the past 3 months? ***TICK ALL THAT APPLY; DO NOT INCLUDE ACTIVITIES THAT ARE PLANNED BUT HAVE NOT YET STARTED.***

Type of channel	**Has taken place in past 3 months (TICK ALL**

	THAT APPLY)
Radio	
--Radio spot	
--Radio coverage by local reporters (such as a news report about your site)	
--Radio call-in talk show	
--Other (radio)	
Television	
--TV spot	
--TV coverage by local news reporters about the MC service	
--TV call-in talk show	
--Other publicity (TV)	
Print and audiovisual media	
Newspaper ad	
Billboard	
Posters (in clinics)	
Posters (in other public places)	
Pamphlet (or printed flyer):	
--For MC client	
--For spouse or partner of client	
--For general population (different from client or spouse pamphlet)	
Video for prospective clients (to show in waiting room)	
Video for general population	
Community-level events:	
Van, truck or other mobile vehicle that circulates in the community to promote MC	

Visits/talks/mobilization in the following venues:	
--Group meetings in the community	
--Schools	
--Factories, industries, mines, plantations	
--Military installations	
--Churches, mosques	
--Beer halls	
--Taxi stands, bus stops, motor bike stands	
--Prisons	
--Meetings with opinion leaders, influentials in the community	
Peer education activities: *(different from mobilization activities above)*	
--Satisfied clients	
Cell phone messages re MC	
Internet website for prospective clients	
Song that promotes male circumcision	
Dramas or plays about MC (such as street theater)	
Testimonials by a celebrity or public figure that has had MC	
Telephone hotline	
Other: ***SPECIFY:***	

That was my last question. Thank you for participating in this interview.

Appendix 3: SYMMACS: INSTRUMENT #1-B: Observation of Male Circumcision Procedures Performed

Instructions: the clinician observes one male circumcision from start to finish. He times the steps in each operation; at the close of the operation and before starting the next observation, he completes this form on the MC procedure observed.

Name of site: ________________________________ City/town and country: ________________

Code for site: ______
Date: __DD __MM __YY
Name/code of clinical observer: ____
Name of the surgical provider(s) performing the MC: ____

Code for the provider observed: ___________________
Cadre of primary surgical provider **performing** the MC:_ physician __clinical officer _nurse
Cadre of secondary surgical provider used to assist in performing/completing the MC (*check all that apply)* ___ physician ___clinical officer ___AMO ___nurse ___other
Cadre of any additional providers assisting primary and or secondary provider during the MC: ___ clinical officer ___nurse ___other

#	**Items to be observed and scored**	**0 = none**	**1= partial**	**2= total**	**Comments**
PREOPERATIVE ASSESSMENT					
	Clinical personnel conduct a basic preoperative assessment including a targeted history and physical exam to exclude surgical contraindications, primarily bleeding disorders, allergies, and immunocompromised states and STIs				
SURGICAL PROCEDURES: INFECTION CONTROL, SAFETY					
	Sterile instruments and consumables used for surgery				
#	**Items to be observed and scored**	**0 = none**	**1= partial**	**2= total**	**Comments**
	Sterile gloves used for surgery				
	Hand washing/disinfection between clients				

	Maintenance of an adequate sterile surgical field when operating				
	Use of protective eyewear by all providers during procedure				
	Safe secure storage and disposal of medical waste by provider/site				
	Correct and hygienic instrument processing				
	Disinfection of surgical beds and areas between patients/clients				
SURGICAL TECHNIQUE: surgeon and/or assisting clinical personnel:					
	Clean surgical area with a recommended surgical scrub solution (chlorhexidine based or Povidine iodine)				
	Correctly identify the skin to be excised				
	Demonstrate “safety first approach’ - ensuring no part of the penis other than the foreskin is in danger of being injured				
	Demonstrate the safe administration of local anesthesia				
	Demonstrate cautious and gentle approach to removing the foreskin				
	Adequately controls bleeding with electrocautery and/or ligating sutures				
	Uses correct technique in tying surgical knots				
	Correctly aligns the frenulum and places secure mattress suture				
	Correctly aligns the other quadrant sutures				
	Avoids placing deep sutures around the frenulum (as the urethra located in the				

	vicinity)				
#	**Items to be observed and scored**	**0 = none**	**1= partial**	**2= total**	**Comments**
#	**Items to be observed and scored**	**0 = none**	**1= partial**	**2= total**	**Comments**
	Places interrupted sutures evenly to avoid leaving gapping margins				
	Ensures no significant bleeding present				
	Places a secure dressing that is not excessively tight.				
POST-OP PROCEDURES AND CARE					
	Staff observe post-op clients for an allergic reaction or any other abnormality before allowing them leave the operating table or recovery room				
	Staff review vital signs				
	Staff provide patients with clear instructions, verbal and written on how to wash and care for the wound, and how to deal with pain and minor bleeding.				
	Staff insist/encourage clients to return for at least one follow up visit or in the case of a complication				
	Staff provide emergency contact details to clients				
	Patients receive post-operative counseling instructions and reinforcement of previous MC/HIV messaging				
	Staff give specific reminders of the 6 week post-operative abstinence period				

TIMING FOR THE PROCEDURE:

Step in the procedure	**Start time**	**End time**

	(minute, second)	(minute, second)
1) Patient enters operating area		
2) Provider scrubs/prepares patient skin (note: applying anesthesia may come first)		
3) Provider administers local anesthesia		
4) Provider removes foreskin *(Start time: 1st incision cut; end time: complete removal of the foreskin)*		
5) Provider performs haemostatis using:		
A. electrocautery **OR**		
B. ligating sutures		
6) Primary provider inserts skin sutures (number of sutures inserted by primary provider = ____)		
7) Secondary provider assists with insertion of skin sutures(Number of sutures inserted by secondary provider =____)*(LEAVE BLANK IF NO SECONDARY PROVIDER)*		
8) Provider applies dressing and cleans the client		
9) Patient leaves operating bed		

Remarks:
____
____

Appendix 4: Efficiency Elements, Number of Procedures, Adverse Effects, and Follow-Up at Each Participating Facility

Code of site manager:	
Date of interview (DDMMYY):	
Code of interviewer:	
Code of facility:	
Type of facility (fixed, outreach, mobile):	
Month/year when adult MCs services began (MMYY):	
Number of service providers that have worked at the MC service in the past week (In total):	
Medical Doctors	
Clinical officers	
Nurses	
Others	
Medical Doctors	

Clinical officers																								
Nurses																								
Others																								
	2010												2011											
<u>EFFICIENCY ELEMENTS</u>	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D
Proportion of operations conducted by:																								
Medical Doctor -%																								
Clinical Officer-%																								
Nurse -%																								
Assistant Medical Officer-%																								
Other - %																								
Surgical technique used:																								
Forceps guided-%																								
Dorsal slit-%																								
Sleeve-%																								

Other-%																								
Which if any tasks does the primary provider share with secondary providers:																								
Surgical preparation																								
Administer anesthesia																								
Suturing for haemostatsis																								
Suturing of skin																								
Haemostatis																								
Bandaging																								
History taking																								
Other																								
Primary provider uses electrocautery (diathermy):																								
Always																								
Sometimes																								

Never																								
Not available																								
Number of surgical beds available:																								
In total																								
Per team																								
Preparation of MC instruments and consumables: % distribution per month																								
Prepackaged consumables / reusable instruments																								
Prepackaged consumables / disposable instruments																								
Nonpackaged consumables / reusable instruments																								
Kits with MC devices																								
<u>NUMBER OF PROCEDURES</u>																								
Number of operations performed per month:																								

ADVERSE EVENTS																								
Nature of adverse event:																								
Intra-operative:																								
Number of bleeding-related AE																								
Other																								
Post-operative:																								
Number of bleeding-related AE																								
Number of infection-related AEs																								
Number of other AEs																								
TOTAL (Intra- and post-operative):																								
Severity of Adverse Events Reported:																								
Severe:																								
Intra-operative																								
Post-operative																								
TOTAL - Severe AE																								

Moderate:																								
Intra-operative																								
Post-operative																								
TOTAL - moderate AE																								
Mild:																								
Intra-operative																								
Post-operative																								
TOTAL - moderate AE																								
TOTAL AE reported per month:																								
<u>FOLLOW-UP</u>																								
Number of follow-up visits conducted																								
Percentage of MC clients returning for follow-up																								
<u>HIV TESTING</u>																								
Percentage of MC clients tested for HIV																								

Appendix 5: Research Approvals

MASENO UNIVERSITY

SCHOOL OF GRADUATE STUDIES

Office of the Dean

Our Ref: PG/PHD/017/2011

Private Bag, MASENO, KENYA
Tel:(057)351 22/351008/351011
FAX: 254-057-351153/351221
Email: sgs@maseno.ac.ke

Date: 28th January, 2014

TO WHOM IT MAY CONCERN

RE: PROPOSAL APPROVAL FOR DICKENS S.O. ADUDA— PG/PHD/017/2011

The above named is registered in the Doctor of Philosophy in Public Health Programme of the School of Public Health and Community Development, Maseno University. This is to confirm that his research proposal titled "Technical Efficiency and Service Quality of Voluntary Medical Circumcision Scale-Up in Nyanza Province, Kenya" has been approved for conduct of research subject to obtaining all other permissions/clearances that may be required beforehand.

Prof. P.O. Owuor
DEAN, SCHOOL OF GRADUATE STUDIES

Maseno University *ISO 9001:2008 Certified*

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840 - 00200 NAIROBI, Kenya
Tel: (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **January 17, 2011**

TO: **DR. N. MURAGURI, PRINCIPAL INVESTIGATOR**
THE DIRECTOR, NASCOP

RE: **NON-SSC PROTOCOL NO. 245 (*INITIAL SUBMISSION*): SYSTEMATIC REVIEW OF MALE CIRCUMCISION SCALE-UP IN KENYA.**

Make reference to your letter dated November 15, 2010 received on November 29, 2010. Thank you for your response to the issues raised by the Committee. This is to inform you that the issues raised during the 183rd meeting of the KEMRI/ERC meeting held on 19th October 2010, have been adequately addressed.

Due consideration has been given to ethical issues and the study is hereby granted approval for implementation effective this **17th day of January 2011,** for a period of twelve (12) months.

Please note that authorization to conduct this study will automatically expire on **16th January 2012.** If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to the ERC Secretariat by **5th October 2011**.

You are required to submit any amendments to this protocol and other information pertinent to human participation in this study to the ERC prior to initiation. You may embark on the study.

Yours sincerely,

RCKithinji..
Caroline Kithinji,
FOR: SECRETARY,
KEMRI/NATIONAL ETHICS REVIEW COMMITTEE

In Search of Better Health

FAMILY HEALTH INTERNATIONAL RECEIVED 02 FEB 2012 P.O. Box 38835 NAIROBI, KENYA

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **January 24, 2012**

TO: **DR. NICHOLAS MURAGURI,**
HEAD, NATIONAL AIDS & STI CONTROL PROGRAM,
PRINCIPAL INVESTIGATOR

RE: **NON SSC NO. 245 (*REQUEST FOR ANNUAL RENEWAL*): SYSTEMATIC REVIEW OF MALE CIRCUMCISION SCALE-UP IN KENYA.**

This is to inform that during the 197th meeting of the KEMRI/ERC held on the 24th of January 2012, the Committee **conducted the annual review and approved** the above referenced application for another year.

The Committee notes that future plans are to:
1. Commence data collection in mid ensuring quality and smooth logistics
2. Ensure coordinated data processing and analysis
3. Attend data analysis workshop
4. Keep all the relevant stakeholders appraised on project progress

This approval is valid from today **January 24, 2012** through to **January 23, 2013.** Please note that authorization to conduct this study will automatically expire on **January 23, 2013**. If you plan to continue with data collection or analysis beyond this date please submit an application for continuing approval to the **ERC** secretariat by **November 1, 2012.**

You are required to submit any amendments to this protocol and other information pertinent to human participation in this study to the SSC and ERC for review prior to initiation.
Yours sincerely,

RCKithinji

Caroline Kithinji,
FOR: Secretary
KEMRI/ETHICS REVEIW COMMITTEE

In Search of Better Health

RECEIVED
07 MAY 2013
FAMILY HEALTH INTERNATIONAL
P.O. Box 38835
NAIROBI, KENYA

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **23rd April, 2013**

TO: **DR. PETER CHERUTICH (PRINCIPAL INVESTIGATOR),
DEPUTY HEAD, NATIONAL AIDS & STI CONTROL PROGRAM (NASCOP),
P.O.BOX 19361-00202
NAIROBI**

Dear Sir,

RE: NON SSC PROTOCOL 245 (*CONTINUING REVIEW REPORT*): SYSTEMATIC REVIEW OF MALE CIRCUMCISION SCALE-UP IN KENYA.

Thank you for the Continuing Review Report for the period of **April 2012- December 2012.** This is to inform you that the request for continuation approval has been reviewed by the ERC.

The Review Team noted that:

(a) Data collection was completed in November 2012
(b) Data analysis is complete
(c) The planned activities for the next project period:
  i. Manuscript writing

The KEMRI Ethics Review Committee was of the considered opinion that the progress made in the reporting period is satisfactory and therefore grants approval for continuation with the study. This approval is effective from **23rd April 2013** for a period of one year.

Please note that the authorization to conduct this study will automatically expire on **22nd April 2014.** If you plan to continue with the study beyond this date, please submit an application for continuation approval to the ERC secretariat by **12th March 2014.**

Yours sincerely,

EAB

**DR. ELIZABETH BUKUSI,
ACTING SECRETARY,
<u>KEMRI ETHICS REVIEW COMMITTEE</u>**

In Search of Better Health

Appendix 6: Participant Information and Consent Form

PARTICIPANT INFORMATION AND CONSENT FORM

for enrollment in the systematic

REVIEW OF MALE CIRCUMCISION SCALE-UP (SYMMACS) PROJECT IN KENYA

(To be administered in English)

Principal Investigators:

1. Dr Nicholas Muraguri, Director, National AIDS & STI Control Program (NASCOP), Nairobi, Kenya
2. Dr Peter Cherutich, Head of HIV Prevention at NASCOP, Nairobi, Kenya

Co-Investigators:

3. Dr Jane Bertrand, Professor and Chair, Department of Health Systems Management

Tulane University School of Public Health and Tropical Medicine, New Orleans, LA. USA

4. Dr Dino Rech, Executive Director, Centre for HIV and AIDS Prevention Studies (CHAPS), Johannesburg, South Africa
5. Dr Kawango Agot, Program Director, Integrated HIV Prevention Interventions including Voluntary Medical Male Circumcision in Nyanza, Kisumu, Kenya
6. Dr Walter Obiero, Clinical Manager, Male Circumcision Project of the Nyanza Reproductive Health Society in Kisumu, Kenya
7. Dr Mores Loolpapit, Associate Director, Male Circumcision Consortium Project of FHI, Nairobi, Kenya

Performance sites: selected health facilities in Nyanza Province, Kenya

Sponsor: USAID through the Research to Prevention (R2P) Project, Johns Hopkins University

Introduction

You are invited to participate in a research study on the scale-up of adult male circumcision services in Kenya. You are being asked to participate because you are a service provider in this male circumcision facility selected for inclusion in the study.

Why is this study being done?

The purpose of this research is to monitor and study the scale-up of male circumcision services in this country and three others (South Africa, Tanzania and Zimbabwe). Specifically, it will track the adoption of certain procedures and means of making service delivery more efficient. The results will help individual clinics identify areas for improvement, and it will help the national program track progress in this effort.

What are the study procedures? What will I be asked to do?

If you agree to take part in this study, the research team will:

1. Observe and time up to five of the male circumcision operations that you perform over the course of your shift today; and

2. Interview you regarding your experiences to date with the provision of male circumcision services, as well as your attitudes and opinions on this topic.

As the letter we sent in advance described, we plan to observe at least 10 operations at this facility today and tomorrow, or at least 5 per surgical provider (whichever number is larger). We will be timing each step in the procedure to learn how long each takes. Also, we will be observing the conditions of the clinic, using questions from the WHO Quality Assurance protocol. After you have finished operating for the day, we would like to interview you to learn about your training in male circumcision, the surgical methods you use, your attitudes and opinions related to the scale-up of male circumcision services in Kenya and related topics. This interview will take approximately 25 minutes. We plan to interview every surgical provider and nurse involved in male circumcision services in a sample of MC facilities.

What are the risks or inconveniences of the study?
We realize that our presence in the operating area may interrupt the patient flow, but we are here to observe what happens without interfering in the normal routine. There is no known risk to your participating in this study. We are not here to supervise your work. Should we observe any areas in which we believe the clinic might want to improve, we will discuss these openly and constructively at the end of our two-day visit to this facility. We realize that you have a very busy schedule and that the interview may be an inconvenience.

What are the benefits of the study?
This study may not benefit you directly. However, it will help us to learn more about the different methods that male circumcision providers and clinics are using to become more efficient in the delivery of this service, while still maintaining safety and quality. The results of this study are expected to guide the design of future programs and to influence other clinics to adopt more efficient methods, with the end result of providing greater numbers of boys and men with the benefits of male circumcision.

Will I receive payment for participation?
You will not be paid to be in this study.

Are there costs to participate?
There are no costs to you to participate in this study.

How will my personal information be protected?
Although we have recorded your name on this questionnaire to ensure that we have interviewed all relevant staff at the clinic, we will remove all names once we transfer this information to the computer. Your name will not appear in any report, nor will we use any description that could identify you (such as “the health care provider at Clinic X”).

The research team will prepare the results in different formats: data on individual facilities, on all the selected facilities in Nyanza Province, and on all participating clinics from the four countries. The following procedures will be used to protect the confidentiality of your data. The researchers will keep all study records (including any codes to your data) locked in a secure location. Research records will be labeled with a unique code. Only the members of the research staff will have access to the passwords. Data that will be shared with others will be coded as described above to help protect your identity. At the end of this study, the researchers may publish their findings. Information will be presented in summary format and no individual will be identifiable in any publications or presentations. Any master key and other data described in this paragraph will be maintained in accordance with the security provisions of this paragraph until destroyed by the researchers.

Can I stop being in the study and what are my rights?
You do not have to be in this study if you do not want to. If you agree, but later change your mind, you may drop out at any time. There are no penalties or consequences of any kind if you decide that you do not want to participate. Also, in the interview you do not have to answer any question that you do not want to answer.

Who do I contact if I have questions about the study?
Take as much time as you like before you make a decision to participate in this study. We will be happy to answer any question you have about this study. If you have further questions about this study, want to voice concerns or complaints about the research or if you have a research-related problem, you may contact the country coordinator for this study, **Dr Mores Loolpapit** at **+ 254 722 227 404** or **+ 254 20 2824 000**.

If you would like to discuss your rights as a research participant, discuss problems, concerns, and questions; obtain information; or offer input with an informed individual who is unaffiliated with the specific research, you may contact the secretary, KEMRI/ National Ethics Review Committee at + 254 20 2722541 or by email at director@kemri.org and also the Tulane University Human Research Protection Office by email: irbmain@tulane.edu. Both entities have provided ethical approval for this project.

<u>Statement of Consent:</u>
I have read this form and decided that I will participate in the research project described above. Its general purposes, the particulars of involvement and possible risks and inconveniences have been explained to my satisfaction. I understand that I can withdraw at any time. My signature also indicates that I have received a copy of this consent form.

____ ____________

Participant Date

____ ____________

Parent/Legally Authorized Representative (if applicable) Date

____ ____________

Person Obtaining Consent / Interviewer Date

____ ____________

Witness Date

Study: Systematic Review of Male Circumcision Scale-up (Symmacs) in Kenya
Participant Information and Consent Form
Version 2.0, 11/15/10