COMPARISON OF EFFICACY OF SEROLOGICAL AND PHOTOGRAMMETRIC TESTING ALGORITHMS FOR HIV SCREENING IN NONLABORATORY SETTINGS

EDWARD IRERI MUGAMBI

MASTER OF SCIENCE

(Medical Virology)

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

2023

Comparison of Efficacy of Serological and Photogrammetric Testing Algorithms for HIV Screening in Nonlaboratory Settings

Edward Mugambi Ireri

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of Master of Science in Medical Virology of Jomo Kenyatta University of Agriculture and Technology

2023

DECLARATION

This thesis is my original work and has not been presented for a degree in any other university.

Signature…….Date……

Edward Mugambi Ireri

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

Signature…….Date……

Prof. Marion W. Mutugi, PhD

JKUAT, Kenya

Signature…….Date……

Prof. Kiptoo Michael, PhD

KEMRI, Kenya

DEDICATION

I would like to dedicate this project to my parents Mr Seraphine Ireri and Mrs Consolata Ireri for their selfless love, encouragement and concerted effort to ensure that I never lacked as I pursued education. I also dedicate this work to my wife Beatrice for believing in me and cheering me on even when the hurdles in my academic journey seemed insurmountable. To my daughter, Princess Maria, and my sons Richie Ireri and Yves Saint Laurent, may you grow to scale the ladders of academia to the highest levels humanely possible.

ACKNOWLEDGEMENT

It has been a long journey that has finally come to an end. In the course of this journey, I have received a lot of support, help and encouragement. Without the assistance of many individuals, I would never have completed this thesis and my MSc study. Therefore, I would like to express my earnest appreciation. Firstly, I would like to express my gratitude to Prof. Marion Mutugi, Mr Lawrence Muthami, and Prof. Michael Kiptoo for the great mentorship and guidance during my long study. Secondly, I am also very thankful to all the HIV Testing and Counseling (HTC) operators within the county of Nairobi who accepted to participate in this study. Not forgetting the entire staff at the National HIV Reference Laboratory (NHRL) in Nairobi, led by Mr Mamo Umaro, and Mrs Nancy Bowen for their cooperation. Thirdly, I would like to thank Mr Abdullah from Saudi Arabia for his enormous financial support towards my MSc study. Lastly, special gratitude goes to the Lord Almighty for giving me strength and good health as I pursued my studies.

God bless you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS AND ACRONYMS

AIDS	Acquired immunodeficiency syndrome
AUC	The area under the curves
CDC	Centre for Disease Control
DASCO	District Aids Committee
DSRS	Department of standards and regulatory services
DTS	Dried tube specimens
DNA	Deoxyribonucleic Acid
DVD	Digital Versatile Disc
ELISA	Enzyme-linked immunosorbent assay
ERC	Ethical Review Committee
FPR	false positive rate
HBTC	Home-Based Testing & Counselling
HIV	Human Immunodeficiency Virus
HTC	HIV Testing and Counselling
ICT	Information Communications and Technology
KEMRI	Kenya Medical Research Institute
LAB	Laboratory
MFL	Master Facility List
MOH	Ministry of Health
NACC	National AIDS Control Unit
NASCOP	National AIDS and STI Control Program
NBTC	National Blood and Transfusion Centre
NHRL	National HIV Reference Laboratory Services

NPHLS	National Public Health and Laboratory Services
PITC	Provider-Initiated HIV Testing and Counselling
PMTCT	Prevention of Mother-to-Child Transmission
PT	Proficiency Testing
RNA	Ribonucleic Acid
ROC	Receiver Operating Characteristic
SPSS	Statistical Program for Social Sciences
SOPs	Standard Operation Procedures
SSC	Scientific Steering Committee
Sw	within-subject standard deviation
TPR	True Positive Rate
TQM	Total Quality Management
VCT	Voluntary Counselling and Testing
WBS	Whole Blood Sample
WHO	World Health Organization

ABSTRACT

Proficiency testing for Human Immunodeficiency Virus offers a platform for institutions and individuals to assess their technical competencies. It boosts morale and confidence among the laboratory operators and provides opportunities for improvement among participating institutions. Post-market field validation of HIV rapid diagnostic test is important because it increases confidence in the quality of testing services. Moreover, it confirms that the kit prequalified by WHO, is of superior quality (and vice versa) in terms of sensitivity and turnaround time. However, a shortcoming in test validation is that once kits are procured, rarely does field validation take place. As of 2008, no study in East Africa in general and Kenya, in particular, had looked at the field validation of photogrammetric testing. This current study is thus aimed at comparing the efficacy of serological and photogrammetric testing for HIV screening in non-laboratory resource-limited settings within Nairobi County. Besides, the study explores a new cost-effective method which can complement routine proficiency testing as a training tool for field-based validation. A longitudinal study was conducted using three rounds of proficiency and photogrammetric testing respectively. A total of 234 experienced and non-experienced operators were recruited using purposive sampling. A total of 702 dried tube specimen panel samples were tested using the Determine algorithm, while 702 photos were visually interpreted. The study revealed that the validity of serological testing was 98.07% and 96.21% for sensitivity and positive predictive values respectively and 70.37% and 82.61% for specificity and negative predictive values respectively. The validity of photogrammetric testing was 96.43% and 98.63% for sensitivity and positive predictive values respectively, and 70% and 46.7% for specificity and negative predictive value respectively. Lastly, the overall accuracy was 94.5% and 95.30% for proficiency and photogrammetric testing respectively, calculated as the percentage of true positives and true negatives on overall results. This study concludes that Determine algorithm is still sensitive and specific as such, it can still be used for proficiency testing of HIV panels. Besides, it was established that photogrammetric testing could be interpreted with higher accuracy compared to proficiency testing. The study recommendes that a higher accuracy rate of interpretation of HIV photographed tests can be used to assess proficiency levels of operators, while photogrammetric testing is viable as a complementary tool for identifying and monitoring operators testing competencies in resource-limited settings.

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Validation of test kits involves field evaluation and surveillance to monitor performance in non-ideal field conditions. Surveillance involves pre-market evaluation while field evaluation involves post-market surveillance or batch-to-batch validation. Therefore, validation can be conducted on a test kit, a test method or an actual test done by service providers. All the three methods that are employed in conducting validation have their merits and demerits. For example, test kit validation has challenges namely: the transportation and storage conditions are likely to expose reagents and test kits to adverse environmental conditions which in turn might negatively influence the test results. At the time of this study, Determine algorithm was the first-line test in Kenya. Thus, post-market field validation of Determine algorithm was deemed important because it increased confidence in the quality of testing services and confirmed that the kit prequalified by WHO is of superior quality (and vice versa) in terms of sensitivity and turnaround time.

Proficiency testing (PT) was launched in Kenya in 2007. A study by Mwangi *et al.*, (2012) reported that innovative approaches were important in HIV Testing and Counseling (HTC), especially in mainstreaming quality, given the complexity of staffing and retention of the right personnel and training of operators. Thus, this current study aims at evaluating photogrammetric testing as a complementary tool for proficiency testing. Photogrammetric testing plays a significant role in identifying operators who require refresher HTC training courses. This is because staff play a key role in the provision of quality services the reason why it is important to conduct a field-based evaluation of the proposed new method (photogrammetric testing). An HTC Innovation, like any other, has the potential to cut down the costs of implementing proficiency testing and increase the number of operators participating in the scheme.

The prevalence rate of HIV in Kenya dropped from 5.9% in 2016 to 4.8% in 2021. This is partly attributed to many programs that are focused on HIV reduction, and voluntary counselling and testing in both rural and urban areas. Additionally, the report indicated that the prevalence rates in females and males is 5.2% and 4.5% respectively (NACC, 2021). The report further stated that Kenya has made tremendous efforts towards alleviating the HIV epidemic resulting in a 68.5% reduction in new infections between 2013-2021. Among the counties that met global and national targets of reducing new HIV infections by 75% between 2013 and 2021 are Murang'a, Nyamira, Turkana, Nyandarua, Homa Bay, Siaya, Nyeri, Migori, Samburu, Kirinyaga, West Pokot, Kisii, Garissa, and Kiambu (NACC, 2021). However, it is worth noting that five counties namely: Nairobi, Kakamega, Bungoma, Vihiga and Busia, recorded increased HIV infections.

According to the National AIDS Control Council report of 2021, 1.43 million adult Kenyans and 78,465 children between the ages of 0-14 years, were living with HIV. The prevalence rate of HIV infections as of 2020 was 4.3% (2.9% in males and 5.5% in females). Among female sex workers, the prevalence rate of HIV infections in 2020 was 29.3%. Moreover, among men who have sex with men and among people who inject themselves with drugs the prevalence rates of HIV infections were 18.2% and 18.7% respectively.

It is equally important to note that Kenya registered a 43% reduction in AIDS-related deaths among adolescents and young people between the 2015 and 2021 period. Overall, the treatment coverage increased from 78% in 2013 to 86% in 2020. Nevertheless, the segregated data by gender showed that the treatment coverage for males above 15 years of age declined from 80% to 77%, while that of females in the same cohort increased from 77% to 91%. With regards to babies, the treatment coverage increased from 42% to 84% (NACC, 2021). The overall AIDS-related deaths reduced from 58,465 in 2013 to 19,465 in 2020. Among males aged 15 years and above, the deaths declined from 20,765 in 2013 to 8,885 in 2020. Among females aged 15 years and above, the figures declined from 27,310 in 2013 to 7,508 in 2020. The AIDS-related deaths among babies declined from 10,390 in 2013 to 3,092 in 2020. In 2020, new HIV infections constituted 61% among adolescents and

younger adults between the ages of 15-29 years old with Nairobi, Homa Bay, Uasin Gishu and Meru counties having the highest number of new HIV infections among young people. Nevertheless, the proportion of new HIV infections among young people aged 15-24 years declined by 69% from 35,776 in 2015 to 11,229 in 2021.

A decrease in HIV-related stigma and innovative testing strategies led to more Kenyans being aware of their HIV status in comparison to the onset of the epidemic. The adoption of testing innovations that are availed locally in community-based facilities improved efficiency and enabled individuals who experienced difficulties assessing facility-based testing, such as VCT, know their HIV status. Community-based testing include mobile testing in the community, a one-stop model, work testing, partner and family testing, social network testing, home-based testing and HIV self-testing. Specific testing methods that have been used in Kenya include HIV rapid diagnostic test kits (RDTs), Enzyme-Linked Immunosorbent assay (ELIZA), polymerase chain reaction (PCR) and Western blot. However, with improvement in technology, new HIV testing technologies have been or are being tested. Some of those that have been tested include machine learning algorithmic solutions that use artificial intelligence and which incorporate ideas from computational learning theory, artificial neural networks, statistics, stochastic modelling, genetic algorithms and pattern recognition (Dubey, 2016). Nevertheless, such technologies have not yet been put to use in Kenya.

HIV testing in Kenya has experienced a myriad of challenges. For example, there have been systemic challenges related to procurement and supply chain management of test kits. Another major challenge is the fear of visiting testing facilities for the purposes of knowing one's HIV status by a section of the populace (NACC, 2021). This group of individuals fear testing because should they turn positive they are afraid of getting stigmatized by their communities (Bott *et al.*, 2015).

1.2 Statement of the problem

In 2011, the standard diagnostic Bioline HIV testing kits were recalled in Kenya after failing the WHO prequalification assessment. The kit had diagnosed patients as HIV negative when they were positive. This finding inspired the study to evaluate and

validate the field performance of Determine test kit in a non-laboratory setting. It is worth noting that at times depending on the number of clients visiting a testing facility, some test kit stocks can take longer periods in the field sites before they are all used up and new stocks are received a factor, which can cause variations in test results. Another shortcoming in test validation is that ones kits are procured, rarely does field validation take place. Besides, HIV rapid test kits used in developing countries are limited to test kits technical performance which only focus on sensitivity, specificity and predictive values without paying much attention to the ease of performance of inter and intrapersonal evaluation. As of 2008, no study in East Africa in general and Kenya in particular, had looked at the field validation of photogrammetric testing. Kate *et al.* (2008), in a study conducted in laboratory settings, focused on the validation of HIV photo results of Determine algorithm. In this study, only two laboratories from Kenya participated and the country sent one set of interpretations on photographed results of rapid HIV assays. Routine proficiency testing scheme was implemented in Kenya since 2007 to monitor and improve the quality of HIV Testing and Counselling services. Nonetheless, its implementation was marred by numerous challenges like logistics, variability in results interpretations by operators and high costs among others. An HIV proficiency testing study conducted in Kenya by Muchiri *et al.* (2016) showed the operators were able to correctly identify 89% of DTS panels whereas 11% had a detection error. This finding revealed that the operators were not fully capable of identifying positive HIV RDTs. It is therefore against this backdrop that the study explores, evaluates, and validates photogrammetric testing in non-laboratory resource-limited settings as a new cost-effective method that complements the routine proficiency testing in terms of field-based validation of HIV rapid algorithms.

1.3 Justification of the study

This study was informed by the presence of limited data on field evaluation of Determine test kits in Kenya regarding specificity, sensitivity, positive predictive value, negative predictive value and accuracy. It is particularly important to evaluate the performance of the assays through field validation and market surveillance

because performance is subject to variation. In Kenya, the National AIDS Control Council (NACC) carries out both initial and post-market validations for HIV test kits. Nevertheless, the data on the validation evaluation is not accessible to the public. Thus, the need for other independent research institutions to validate the performance of the assays in the market. In Mwangi *et al.* (2012) report, logistics management of HTC services, the high demand for HIV services, weak systems and financial implications, are some of the challenges that need innovative approaches. Therefore, due to the many problems associated with implementing the PT schemes in Kenya, there is a need to explore an integrative and innovative approach (photogrammetric testing) to complement the existing scheme. This innovative approach will help the National HIV Reference Laboratory to identify operators that require urgent refresher training in HTC and cut down on the cost of conducting PT on every HTC operator. This study thus fills the innovation gap as it compares the efficacy of serological and photogrammetric testing for HIV screening in non-laboratory resource-limited settings.

1.4 Research questions

1. What is the specificity and sensitivity of HIV proficiency testing panels results in a nonlaboratory-based setting in the field?
2. What is the specificity and sensitivity of HIV photogrammetric testing results in a nonlaboratory-based setting in the field?
3. What is the feasibility of using photographed rapid HIV test results to monitor testing competencies amongst resource-limited HTC sites within Nairobi county?

1.5 Study objectives

1.5.1 General objective

To compare the efficacy of serological and photogrammetric testing algorithms for HIV screening in non-laboratory settings

1.5.2 Specific objectives

1. To determine the specificity and sensitivity of HIV proficiency testing panels results in a nonlaboratory based setting in the field.
2. To determine the specificity and sensitivity of HIV photogrammetric testing results in a nonlaboratory based setting in the field.
3. To assess the use of photographed rapid HIV tests results as a monitoring tool for testing competencies and training operators in nonlaboratory based settings.

CHAPTER TWO

LITERATURE REVIEW

Test kits validation and proficiency testing are processes of external quality assessment schemes. According to WHO (2009), HIV testing should be aligned to national algorithms, including the use of HIV assays approved and validated by the national reference laboratory. Ongoing quality assurance is required to monitor and evaluate the performance of each test within the national algorithms, to ensure successful performance of the testing technology and algorithm.

There exists a very extensive literature on the numerous assays that have been developed for HIV antibody detection and promotion of HIV screening and diagnosis. For example, the simple rapid test kits are popular especially for field testing in resource-limited settings because they are easy to use since no instrument is required; they are cost-effective and they do not require cold chain storage (Lyamuya *et al.*, 2009) (Boadu *et al.*, 2016). Moreover, they are single-use, disposable devices that may be used to directly test whole blood specimens, serum, plasma, and/or oral fluids (WHO, 2005). HIV intervention programs such as Prevention of Mother- to- Child Transmission (PMTC), Home Based HIV Testing & Counselling (HBTC), Provider Initiated Testing & Counselling (PITC), and Voluntary Counselling & Testing (VCT) require the use of simple rapid assays because they give same-day results. According to WHO (2005), performance characteristics and operational requirements of an HIV assay should promote and increase access to HIV testing services, especially in resource-limited countries.

Rapid HIV algorithms were designed to help providers offer counselling and testing services quickly and easily. They are advantageous because they increase preventive measures (VCT) and interventions (PMTCT); they support increased number of testing sites; they offer same-day diagnosis and provide feedback immediately; they are robust and their test time is under 30 minutes; most of them do not need refrigeration or reagents; they require minimal technical skills as well as minimal equipment or none at all among other merits. Since rapid HIV testing in resource-limited settings is supposed to be simple, it must be performed with utmost accuracy

by both trained and untrained operators in accordance to the instructions on the algorithm insert pack. Thus, according to Mandrekar (2010b) when describing a diagnostic test, it is important to report both sensitivity and specificity because they are inherently linked. That is, as the value of one increases, that of another decreases. Special attention must be paid to the financial and emotional implications of a disease because higher sensitivity is often considered desirable in a diagnostic setting, while higher specificity is desirable in a screening set up. Sensitivity and specificity are measures of intrinsic diagnostic accuracy because they are not affected by the prevalence of a disease condition (Zhou *et al.*, 2009).

An HIV assay evaluation study by Kroidl *et al.* (2012) found that the specificity of Determine algorithm was highly dependent on the tested sample type. The acceptability of HIV rapid tests is generally high in medical settings but lower in non-medical settings. This observation is attributed to the fact that in medical settings, the experimental conditions are well controlled to replicate almost similar conditions that the manufacturers had validated in the algorithms. HIV rapid algorithms also pose some problems especially with regards to sensitivity or when they are used in a context different from the laboratory (Ha *et al.*, 2014; Roberts *et al.*, 2007). For example, a study conducted by Black *et al.* (2009) reported that three routinely used kits, which had previously performed well under laboratory conditions, had lower sensitivity rates under clinical settings. Additionally , HIV rapid test kits were reported to have underperformed in field validation thereby failing to detect a substantial number of infections. Besides, there were sensitivity differences between nurses and laboratory technicians at varied antenatal sites (Moodley *et al.*, 2008). The positive and negative predictive rates of the rapid diagnostic algorithms strongly influence the test performance in a population (Boadu *et al.*, 2016). This means that the sensitivity and specificity results from algorithms evaluation studies before licensing and marketing are not necessarily achieved in the field. Besides, Abokyi *et al.* (2014) reported cross-reactivity when serum is used to prepare HIV panels. The report by WHO (2004) noted that rapid tests are useful in resource-limited settings although their poor negative predictive values can be a cause of alarm especially when used in populations with high HIV incidences and prevalences. Consequently, when performed by inexperienced personnel, Rapid tests

may post results that are flawed thereby giving incorrect positive and negative outcomes (Black *et al.*, 2009).

Previous studies have largely focused on validation of proficiency testing using Determine algorithm under laboratory settings (Koblavi-Dème *et al.* (2001); Urassa *et al.* (2002); Van den Berk *et al.* (2003); Rouet *et al.* (2004); Tegbaru *et al.* (2004); Granade *et al.* (2005); Singer *et al.* (2005); Gray *et al.* (2007); Eller *et al.* (2007); Mayhood *et al.* (2008); Anzala *et al.* (2008); Lyamuya *et al.* (2009); Piwowar-Manning *et al.* (2010); Zeh *et al.* (2011); Kroidl *et al.* (2012) and Chakrabarty *et al.* (2015). However, a limited number of studies have been conducted to show how Photographed HIV test results can be used to provide a novel, cost-effective approach to EQAS for non-laboratory rapid HIV testing. Once such approaches are established, programs could be used for training and monitoring purposes, facilitating the more accurate interpretation of rapid HIV assays thereby assisting in HIV prevention efforts in resource-limited countries (Kate *et al.*, 2008). The following section will review literature and identify gaps as it compares the efficacy of serological and photogrammetric testing for HIV screening in non-laboratory resource-limited settings.

2.1 Deep learning application and HIV proficiency testing

Deep learning applications are performed using python and R programming. The deep learning algorithms are known to harness the advances made in large datasets while at the same time processing power. The performance abilities of the deep learning applications have been shown to exceed human performance in tests that require visual tasks (acuity) (De Fauw *et al.*, 2018; Doan & Carpenter, 2019; Esteva *et al.*, 2017; Xu *et al.*, 2019) when used alongside appropriate guidelines (Ching *et al.*, 2018; Rajkomar *et al.*, 2019). On a similar note, some studies such as those conducted by (Carrio *et al.*, 2015; Zeng *et al.*, 2016) have also been investigating the application of deep learning to the interpretation of Rapid Diagnostic Tests (RDTs). However, it is worth noting that the deep learning application of RDTs is dependent on the uniformity of images, harsh environmental factors and a variety of test types.

Technology-based deep learning application has undergone proficiency testing and has shown promising results compared to the traditional visual interpretation by humans. In a research conducted in rural South Africa, a pilot field study of the deep learning algorithm was deployed as a mobile application and its findings demonstrated high levels of sensitivity (97.8%) and specificity (100%) compared to the traditional visual interpretation by humans (Turbé *et al.*, 2021). The findings laid the foundations for a new paradigm of deep learning-enabled diagnostics in low-and middle-income countries. The findings also made provisions for the following: a platform for workforce training, quality assurance, decision support and mobile connectivity to inform disease control strategies, strengthening healthcare system efficiency and improving patients outcomes and outbreak management in emerging infections (Turbé *et al.*, 2021).

2.2 Machine learning and HIV testing

Machine learning is a branch of artificial intelligence which revolves around the design and development of algorithms that allow computers to evolve behaviours based on empirical data. Machine learning recognizes complex patterns automatically and makes intelligent decisions based on training and test data. The training data is the one set aside to develop a model that will be used to test data to find out if it fits. In most cases, the training data comprises 75% of the entire data, while the test data takes up the remaining 25% (Dubey, 2016).

The field of machine learning has attracted attention in the field of HIV diagnosis, screening, treatment, design and production of vaccines for the cure of HIV. With its algorithmic solutions, machine learning incorporates ideas from computational learning theory, artificial neural networks, statistics, stochastic modelling, genetic algorithms and pattern recognition (Dubey, 2016). The study further notes that machine learning methods are fast and they can economically complement wet laboratory techniques since they provide methods, techniques and tools that help solve diagnostic and prognostic problems in a variety of medical domains.

2.3 Test Kit Validation

Laboratory and field performances of many HIV Rapid tests have been evaluated in the programmes of many countries with excellent results being posted. These evaluations have been done using both parallel and serial algorithms coupled with a tiebreaker. In a paper presented at the National research consultative meeting in Nakuru, Kenya, Tukei *et al* (2004) found that the rates of discordancy were low <0.5 per cent and serial testing though cheaper, was not appropriate with HIV prevalence rates above 10%. The study concluded that rapid tests are offering a unique opportunity to speedily expand HIV testing programmes to include rural and hitherto neglected areas.

The field or laboratory appraisal performance of HIV assays mainly rely on the investigation of the following: sensitivity rate, specificity rate, the positive predictive value (PPV), the negative predictive value (NPV) besides investigating the area under the curve (AUC) and the receiver operative characteristics (ROC). The performance of HIV rapid diagnostic testing in a population is influenced by positive and negative predictive values of the rapid diagnostic test kits used (Boadu *et al.*, 2016). On the contrary, the sensitivity and specificity results from test kits evaluations studies before licensing and marketing of the kits is not necessarily achieved in the field. This is because each RDT kit has been designed to detect HIV antibodies either in whole blood, serum, or saliva and each sample would react differently when using a particular kit. For example, in study conducted in Ghana Abokyi *et al.* (2014) found that the lower specificity value of first response HIV-1-2 RDT kit was as a result of cross-reactivity when serum was used as the test specimen. This meant that RDT kit may pose a challenge with serum specimen as compared to the whole blood specimen. In their evaluation study, Jaspard *et al.* (2014) compared finger stick blood (FSB) specimens with oral fluid (OF) specimens. The study revealed unexpected differences in performance with variations in sensitivity and specificity, where FSB test showed greater reliability as compared to the OF tests.

Determine is an in vitro, visually read, qualitative immunoassay. It is used for the detection of HIV-1 p24 antigen and antibodies to HIV-1 and HIV-2 in human serum, plasma, capillary whole blood, or venipuncture whole blood. Determine™ test kit is rapid, simple to use and useful when it comes to the detection of HIV antibodies (Singer *et al.*, 2005; Van den Berk *et al.*, 2003). A limited number of comparative studies on the sensitivity, specificity, the negative and positive predictive values of Determine test kit under limited-resource settings are available especially in Africa. In the literature that follows, the focus will mainly be on studies that evaluated and validated Determine test kit either in a laboratory or field-based setting.

2.4 Sensitivity

According to Mandrekar (2010b), sensitivity is the probability of a test to detect a disease when it is truly present. It is given by the ratio of true positive tests (true positives + false negatives). According to Buttò *et al.* (2010), screening assays are designed to detect all individuals who are infected therefore, they must have a high degree of sensitivity (low false-negative rate). Confirmatory assays on the other hand, must have a high degree of specificity (low false-positive rate). Determine™ HIV-1/2 Ag/Ab Combo test kit insert pack reported a clinical performance sensitivity of 99.9% using serum, plasma, venous whole blood, and capillary (finger stick) whole blood. Similarly, studies done on sensitivity using whole blood samples reported sensitivity rates of 100%. For example, the field evaluation studies in Tanzania by Lyamuya *et al.* (2009) and Chakrabarty *et al.* (2015) ; the one conducted in Uganda by Eller *et al.* (2007); the one done in Cameroon by Granade *et al.* (2005); the ones conducted in Kenya by Anzala *et al.* (2008) and Zeh *et al.* (2011)among others. On the contrary, there are studies done on sensitivity using samples of serum. For example, Koblavi-Dème *et al.* (2001) and Rouet *et al.* (2004) studies in Ivory Coast and the one done in Uganda by Singer *et al.* (2005) among others. All these studies just like those done using whole blood samples reported sensitivity rates of 100%. Other studies reported sensitivity rates lower than 100%. For example, the study conducted by Tegbaru *et al.* (2004) using both whole blood and plasma samples reported 99%; the one by Anzala *et al.* (2008) reported sensitivity rate of 98.98% in Kakira and 97.63% in Masaka both in Uganda, and

98.78% Kangemi in Kenya and the one by Mayhood *et al.* (2008) reported 99.6% among others.These studies give detailed explanations of field evaluation of Determine algorithm using serum, whole blood, and plasma samples.

2.5 Specificity

Specificity is the probability of a test to exclude the disease status of patients who do not have the disease (Mandrekar, 2010b). Determine^{TM} HIV-1/2 Ag/Ab Combo test kit insert pack reported a clinical performance sensitivity of 99.6% using serum and plasma, 99.7% sensitivity using venous whole blood and capillary (finger stick) whole blood (99.8%). Several landmark studies observed specificity slightly lower than 100%. For example, Urassa *et al.* (2002) reported a specificity rate of 97.9%; Lyamuya *et al.* (2009) study in Tanzania and Chakrabarty *et al.* (2015) study in Bangladesh reported 99.6%; Koblavi-Dème *et al.* (2001) study in Ivory Coast reported 99.4%; Rouet *et al.* (2004) study reported 98.4%; and Zeh *et al.* (2011) study in Kenya reported a specificity rate of 99.1% among many others. Extensive research on field evaluation in resource-limited settings still needs to be done. Positive Predictive Values of a new proposed test are very important as they give a clear indication of how good a test is at predicting the status of a disease in a patient (Mandrekar, 2010b).

2.6 Positive Predictive Value

The Positive Predictive Value (PPV) is the probability that a patient has a disease given that the test results are positive, that is, true positives / (true positives + false positives). Studies done on the Positive Predictive Value (PPV) on Determine test kits have been rarely reported. Nonetheless, some studies have documented PPV rates. For example, a study done using whole blood sample by Gray *et al.* (2007) reported a PPV rate of 56.3%; Kroidl *et al.* (2012) study done using plasma reported a PPV rate of 82.6% and 32.9% in whole blood; Mayhood *et al.* (2008) study using whole blood reported a rate of 99.5%; a study conducted by Anzala *et al.* (2008) using whole blood samples reported a PPV rate of 65.71% in Kakira and 45.70% in Masaka both in Uganda while a similar one reported a PPV rate of 95.42% in Kilifi and 97.38% Kangemi both in Kenya among other studies. These findings show that

there are very few published results about the PPV of field evaluation of Determine test kit. Thus, the current study contributes to literature by investigating the PPV rate of field evaluation of Determine test kit.

2.7 Negative Predictive Value

The Negative Predictive Value (NPV) is the probability that a patient does not have a disease given that the test results are indeed negative, that is true negatives / (true negatives + false negatives). Although some attempts have been made to address the Negative Predictive Value (NPV) during field evaluation of Determine test kit, NPV findings are rarely reported. However, some studies reported their findings. For example, Gray *et al.* (2007) conducted a study in Uganda using whole blood sample and reported a NPV rate of 99.7%; Mayhood *et al.* (2008) did one in Tanzania using whole blood sample and reported 99.9% NPV rate. In a study conducted by Anzala *et al.* (2008), using whole blood samples a NPV rate of 99.90% was reported in Kakira and 99.76% in Masaka both in Uganda, while a similar study conducted in Kenya reported a 100% NPV rate in Kilifi and 99.89% Kangemi. From the research conducted, it is evident that the topic on NPV field evaluation of Determine test kit needs further investigation and literature documentation.

2.8 Accuracy

Studies on the findings of the accuracy of the Determine test kit are scanty. According to a study conducted by Anzala *et al.* (2008) in which whole blood samples were used, an accuracy rate of 95.71% was reported in Kakira and 90.49% in Masaka both in Uganda. A similar study conducted in Kenya by the same researchers revealed an accuracy rate of 99.60% in Kilifi and 99.60% Kangemi. Previous studies have neglected the aspect of accuracy of the Determine test kit as revealed by the current study. Moreover, this aspect of research has been overshadowed by the findings on specificity and sensitivity.

2.9 ROC and AUC

The Receiver Operative Curve provides an ideal means of studying observer performance of a diagnostic test. ROC is a plot of the sensitivity versus 1 – specificity of a diagnostic test. It should be noted that the ROC is the average value of sensitivity for a test over all possible values of specificity and vice versa. Therefore, it is of critical importance to plot an overall ROC as it is very useful in the early stages of evaluation of a new diagnostic test (Mandrekar, 2010a). The area under the ROC provides a measure of discrimination and allows investigators to compare the performance of two or more diagnostic tests. An area under the curve (AUC) value of 0.5 suggests no discrimination, 0.7 to 0.8 is considered acceptable, 0.8 to 0.9 is considered excellent, while more than 0.9 is considered outstanding. A value of 0.5 indicates that the curve falls on the diagonal line as such, the diagnostic tests has no discriminatory ability.

2.10 Reliability

Reliability is the degree to which a score is stable and consistent when measured at different times (test-retest reliability), in different ways (parallel-forms and alternate-forms), or with different items within the same scale (internal consistency). Repeatability or test-retest reliability is the variation in measurements taken by a single person or instrument on the same item and under the same conditions. A measurement is said to be repeatable when this variation is smaller than some agreed limit. In research, the term reliability means "repeatability" or "consistency." A measure is considered reliable if it would give us the same result repeatedly (Trochim, 2006).

According to the Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, repeatability conditions include the same measurement procedure, the same observer, the same measuring instrument used under the same conditions and the same location, and repetition over a short period. Bland and Altman (1986) developed repeatability methods.

The repeatability coefficient is a precision measure, which represents the value, below which the absolute difference between two repeated test results is expected to lie within a probability of 95%. The standard deviation under repeatability conditions is part of precision and accuracy.

Therefore, repeatability is expressed quantitatively in terms of the dispersion characteristics of the results. Four indicators are mostly used to determine the test reliability of a clinical laboratory test. Two of these, accuracy, and precision, reflect how well the test method performs on a day-to-day basis in a laboratory. The other two, sensitivity and specificity, deal with how well the test can distinguish the presence of a disease from its absence.

To achieve accuracy and reliability of test results, several conditions have to be met for example, there has to be a qualified tester and the environment where the tests are carried out must be conducive. Moreover, testing kits, equipment that suit the attributes of each institute must be availed. Additionally, standardized equipment management, regular attendance at training programs, efficient quality management and participation in external quality assessment must be adhered to (Wang *et al.*, 2011). Studies conducted in Gabon proved that immunoassays such as particle agglutination assays, rapid tests and western or line blots are scored or read subjectively with the risk of intra- and inter-reader variability, as well as intra- and inter-laboratory variability. Several factors such as the geographical origin and conditions of the blood sample, the environmental conditions at the site of testing, human interpretation of the results, and the inherent qualities of the test can influence the performance of serological tests (Makuwa *et al.*, 2002).

In a study conducted in China by lieu *et al.* (2010), on thirty commercial assays for testing HIV, it was shown that 95% credible and 95% confidence intervals can be used to measure the reliability of the Bayesian and Classical estimates of sensitivity and specificity. Moreover, Positive Predictive Values can also be determined, that is, the narrower the interval, the more accurate the estimates.

The accuracy and reliability of HIV rapid testing are critical for the success of the rapid expansion of prevention and treatment programs in resource-limited countries. Studies conducted in Uganda, Haiti, and Botswana on the quality of HIV rapid testing kits presented unique challenges especially with regards to the tests done in various settings by personnel without formal laboratory training. Hence, the need for the development and implementation of a generic HIV rapid test training package using a systematic approach to standardise training and ensure quality of such tests (Yao *et al.* (2010).

Research conducted by Hamilton *et al.* (2008), established that the rapid scale-up of HIV Counselling and Testing programmes in Kenya had led to quality concerns. Issues raised included potential for abuse within the private, confidential setting of client-initiated voluntary Counselling and Testing. Another issue was that the systems for investigating and deregistering counsellors though developed were not formalized.

2.11 Reproducibility

Reproducibility is the variability of the measurement system caused by differences in operator behaviour. Mathematically, it is the variability of the average values obtained by several operators while measuring the same item. The most used method for computing reproducibility is the range and average method.

Careful clinical follow-up and meticulous laboratory evaluation can determine which particular strategies have the potential to be efficacious. It is only in this way can reproducibility of results between all sites and maintenance of high levels of internal and external quality control be ensured (Gotch *et al.*, 2005).

With more than one technologist within a workflow, type-specific reproducibility can be achieved in a laboratory under routine but highly standardized testing conditions. This has been evidenced by a study conducted in the USA on type-specific reproducibility of the Roche linear array HPV genotyping test (Steinau *et al.*, 2008). This study emphasizes that even though inter-assay comparability studies had been conducted, there is need to undertake studies on the degree of intra-assay

reproducibility. Intra-assay precision and reproducibility study conducted in Australia noted a good correlation between the assays used across the linear range, although their concordance at the clinically critical lower limit of quantification was poor (Yan *et al.,* 2010). The accurate quantification of low-level viremia remains elusive. Moreover, lack of correlation of the assays not only presents a challenge to the interpretation of the results but also in the clinical management of HIV infected patients.

Past evaluation of rapid diagnostic tests mostly concentrated on sensitivity and specificity. However, a proper evaluation of rapid diagnostic test should address its performance (sensitivity, specificity, and reproducibility) as well as its operational characteristics (user-friendliness and stability) and cost (Boelaert *et al.,* 2007). It has been established that there are no published articles on intra-lab operator precision and inter-VCT sites reproducibility studies through laboratory testing competence in Kenya.

2.12 Accuracy and Precision

The accuracy and precision of each test method have been established and are frequently monitored by professional laboratory personnel. Data on sensitivity and specificity, which are derived from research studies, are generally found in medical literature. Laboratory tests are designed to be as precise, accurate, specific, and sensitive as possible. This is despite the fact that each test has its performance measures and appropriate uses. The specifications given in regards to the design of a laboratory test are the cornerstones of the reliability of test results and provide the confidence that a health care provider has in using the clinical laboratory (Lab tests online). Statistical measurements of accuracy and precision reveal a laboratory test's basic reliability. Accuracy and precision, which describe sources of variability, are not interchangeable. A test method can be precise (reliable reproducibility) without being accurate (measuring what it is supposed to measure and its true value) and vice versa. The level of precision and accuracy that can be obtained is specific to individual test methods but is constantly monitored for reliability through comprehensive quality control and quality assurance procedures. Therefore, when

blood is tested more than once by the same laboratory, the test results should not change much unless the condition has changed. Differences may arise in varied laboratories in relation to precision and accuracy as a result of the different analytical instrumentation or methodologies employed. However, the test results are reported with standardized reference intervals specific to a given laboratory.

A test method is said to be accurate (true) when the test value approaches the absolute “true” value of the substance (analyte) being measured. Results from every test performed are compared to known "control specimens" that have undergone multiple evaluations. They are also compared to the "gold" standard for that assay, thus analysed to the best testing standards available. A test method is said to be precise when repeated analyses on the same sample give similar results. When a test method is precise, the amount of random variation is small therefore, the test method can be trusted because results are reliably reproduced time after time (Lab tests online, 2011).

The tests that a medical provider chooses to use to diagnose or monitor a medical condition are based on their inherent ability to distinguish whether one has the condition or not. Depending on the symptoms and medical history, a provider orders tests to confirm a condition (tests with high sensitivity) or tests to rule out the condition (tests with high specificity). In this regard, sensitivity is the ability of a test to correctly identify individuals who have a given disease or condition whereas specificity is the ability of a test to correctly exclude individuals who do not have a given disease or condition. Currently, there are no published articles on accuracy and precision of VCT HIV Testing and Counselling services in Kenya.

CHAPTER THREE

MATERIALS AND METHODS

This chapter discusses the methodological processes undertaken to conduct the research. It is divided into eleven parts namely: study site, study design, study population, sampling design, sample size, inclusion criteria, exclusion criteria, study procedure, data presentation methods, data analysis as well as ethical considerations.

3.1 Study site

The participating study sites were selected from the eight wards in Nairobi county namely: Dagoretti, Pumwani, Kibera, Kasarani, Embakasi, Westlands, Makadara, and Central.

3.2 Study design

Due to the three PT rounds that were to be investigated in the study, a longitudinal (prospective) study design was adopted. To address objectives 1 and 2 on validity testing, blinded panel sera and photos were distributed after every three months for a period of nine months for field evaluation. Objective 3 was achieved by assessing the competencies of HIV testing service counsellors.

3.3 Study population

The study only involved HIV testing service counsellors who were affiliated to HIV testing facilities in Nairobi County which were registered by the National HIV Reference Laboratory.

3.4 Sampling design

Cluster sampling was used to recruit the study participants based on the Master Facility List (MFL) on the eight clusters which were identified according to 2009 DASCO's in Nairobi. The study focused on all active VCT sites which were already registered under the National HIV Reference Laboratory (NHRL) within Nairobi.

The MFL is a reference number issued by the NHRL to all facilities that have registered to undertake its PT scheme.

3.5 Sample size determination

The sample size for the repeatability study was determined by choosing the appropriate sample size to estimate the within-subject standard deviation, s_w. This procedure is described in statistics notes in the *BMJ* (Altman & Bland, 2010).

It was assumed that the within-subject standard deviation was the same throughout the range.

Another assumption was that within the subject, the distribution of observations was normal, to estimate the standard error. Thus, for the estimation of sample size**,** only cases where there were equal numbers of observations on each subject were considered, as it was not appropriate to plan an investigation with unequal numbers.

The precision of estimating s_w depends on both the number of subjects, *n*, and the number of observations per subject, *m*. A convenient way to deal with the dependence of the standard error, and hence the sample size, on the quantity to be estimated was achieved by estimating it to within a 10 per cent fraction of the population value.

Thus, the sample size was determined using the repeatability formula suggested by Altman and Bland (2010), in which the within-subject standard deviation (sw) was estimated. The standard deviation depends on the number of subjects (n) and the number of observations per subject (m), with the width of 95% confidence interval while estimating the sample within some fraction of the population (10%), assuming a power of 90% (10% of the population assumed to commit a type 2 error).

$$1.96\frac{sw}{\sqrt{2n(m-1)}} = 0.10sw$$

Where the Confidence interval at 95% (*z* score) = 1.96, *n* = no. of subjects, *m* = no. of observations, *Sw* = within-subjects *SD*. This equation has two unknown quantities, so there are many combinations of *n* and *m* which can give the required precision.

In this study, an upper limit on the number of observations per operator was set at three (positive sample, negative sample, and inconclusive sample for both proficiency testing and photogrammetric), therefore *m* = 3. Then, the minimum sample size was 64 operators:

$$\frac{1.96}{\sqrt{2n(m-1)}} = 0.10$$

$$n = \frac{1.96^2}{3 \times 0.10^2 \times 2} = 64$$

The number of MFL in Nairobi County as at the time of the study was 234 each site with several operators. Thus, all the sites were given an equal chance to take part in the study because it was mandatory for them to participate in all government-run proficiency testing schemes. A simple random sampling was then conducted in each cluster to identify the operator who would participate in the 3-follow-up study. This led to the study having 234 study operators participating in the study. The probability of committing a type I error is the same as the level of significance (95% = 0.05). Conversely, the likelihood of committing type II error is the same as the power of the test (90% = 0.10). The reason the power of 90% was chosen was to increase the sample size in order to allow for generalisability of the findings since a minimum power of (80% = 0.20) gives a small sample. Increasing the sample size used in a statistical test is crucial because it reduces the beta risk. An acceptable level of beta risk is 10% beyond that the sample size should be increased.

Therefore, to evaluate the serologic testing algorithm for human immunodeficiency virus (HIV) based on Determine rapid assay, a total of 702 dried tube panel sera with known HIV serologic status (positive, negative, and indeterminate) were used. These 702 photos (positive, negative, and indeterminate) were then interpreted visually. A

total of 234 experienced and non-experienced operators were recruited using purposive sampling to evaluate the feasibility of using photogrammetric testing as a complementary tool for proficiency testing.

3.5.1 Inclusion criteria

1. The participants in registered facilities that have MFL numbers;
2. The participants operate in registered facilities within Nairobi county.
3. The participants are registered members of NASCOP.

3.5.2 Exclusion criteria

1. HIV testing service counsellors in facilities without MFL numbers;
2. Operators whose testing facilities were outside Nairobi County;
3. HIV testing service counsellors who are not registered by NASCOP; and
4. Those who declined to participate in the study.

3.6 Study procedure

3.6.1 Proficiency testing panel description

Two pints of Whole Blood Sample (HIV positive and HIV negative) were obtained from the National Blood and Transfusion Centre (NBTC) in Nairobi. Dried Tube Specimens (DTS) were prepared at the National HIV Reference Laboratory (NHRL) according to the method described by Parekh *et al.* (2010). DTS does not require deep freezers to store, can be easily prepared and reconstituted in the field without the requirement of trained experts. The two pints of blood were characterised using Enzyme-linked immunosorbent assay (ELISA). Several PT panels of dry tube samples (DTS panels: one positive, one negative & one inconclusive) were prepared by transferring 20 µL of plasma, premixed with 0.1% (v/v) green dye, into 2 mL Sarstedt tubes. The tubes which were in a biosafety cabinet were dried overnight at room temperature and stored at 4^{o}C. The operators were supposed to rehydrate the DTS before testing. Random samples of DTS were retested at the NHRL for quality control using Determine® test algorithm. The panels were then packed and dispatched to the participating facilities with the required buffer and instructions for

reconstitution. A self-administered questionnaire which captured demographic and HIV rapid testing information of the operators and the type of the testing facility accompanied the panels. The participants were asked to complete the PT testing within 2 weeks of receiving the panels, fill in the results on a PT sheet and submit via email. The outcome was mailed directly to each participant via his/her email account. The DTS were shipped after a three month interval for the nine months of the study.

3.6.2 Preparation and validation of photogrammetric photos

Before the dispatch of the DTS, random samples we re-tested. Photos of confirmed positive, negative, and inconclusive samples were captured using a Samsung Galaxy S5 mobile phone camera. The mobile phone operated on android 4.4.2 and had a screen area of 1080x1920 pixels and a 16MP (2160p) camera. Photogrammetric testing involves relaying information via mobile technology. It is an effective method because it cuts down on costs as it does not require any use of paperwork. Additionally, the photos captured are clear and interpretation and submission of the results are done in a very short time.

3.6.3 Validation of classification of photographed results

Scanned coloured copies of HIV Determine® results were supplied to three laboratory-trained KEMRI medical virologists for validation purposes. They interpreted the results for the three photos to ensure the line patterns, the shapes of control and the test lines were similar to those obtained from the actual samples. It is worthy pointing out that these interpretations were subjective in nature. A photographer with no laboratory experience confirmed that the colour and contrast of the lines matched the actual RDT results. Once confirmed as a true representative of the findings, the coloured photos were simultaneously dispatched using mobile phones to the participants together with the panels.

3.6.4 Scoring the samples

The HIV testing service counsellors were asked to score each sample based on the three parameters of reactivity namely: positive, negative, or inconclusive. The operators were then asked to complete the interpretation of the photos within 2 weeks (the same time for collecting the PT results). The same photos were again dispatched after three months, a process that was repeated once more during the nine months that the study was conducted. The operators were requested to submit their results via e-mail and the feedback was emailed directly to them through the same means.

3.6.5 Data entry method

Data was entered in MS excel, cleaned and counter checked for double entries. Data coding was done and then entered in SPSS version 25 for further analysis.

3.6.6 Data analysis and presentation methods

Data was analysed using SPSS version 25. Descriptive statistics, which was used to analyse preliminary data, described the distribution of scores and gave a record on the number of times a score occured. The percentage of frequencies was computed to show the differences in sizes. Inferential statistics was used to facilitate the generalization of the results from the sample population. Contingency tables produced using cross-tabulations were used to determine the following: specificity, sensitivity, positive predictive values, negative predictive values, and the Chi-square statistics. Chi-square test was used to establish the relationships between the variables which were categorical in nature.

The graphical visualisation of the results for both proficiency testing and photogrammetric were achieved through plotting the Receiver Operating Characteristic (ROC) curve on proficiency testing and photogrammetric testing using SPSS version 25. The true positive rate (TPR, y-axis) was plotted against the false positive rate (FPR, x-axis) and the Area Under Curves (AUC) determined. Several tables were used to present the confusion matrices and a summary of the sample

characteristics. Numerical exploration of bivariate variables was done using Binary Logistics regression. The cleaned data file, the findings from the study, and the final report were uploaded and saved on cloud via Dropbox.

3.6.7 Quality control

Quality control refers to the efforts and procedures that ensure quality and accuracy of the data collected using specific methods. The current study assessed quality control at three levels: the study design, data collection, and analyses. In the design, the protocol was reviewed based on the supervisors' comments and defended at KEMRI and the workshop defense at JKUAT. During data collection, the same DTS samples were distributed to the participating facilities within one week. This ensured that tested samples were prepared from the same blood samples. The photos were validated by three laboratory-trained KEMRI medical virologists through subjective interpretation. Lastly, the analyses computed were compared using specificity, sensitivity, negative predictive and positive predictive values of the Determine® test.

3.7 Ethical considerations

No blood samples were collected from any VCT clients. The project only commenced after approval and the receipt of the necessary letters from the SSC and ERC committees. The SSC number for the current study was ESACIPAC/SSC/9990, while the ERC number was KEMRI/RES/7/3/1. The pints of blood needed for the preparations of the panel sera were obtained from the National Blood Transfusion Centre in consultation with the NHRL. The samples of pints of whole blood were packed in an icebox and taken to NHRL for storage and further analysis.

CHAPTER FOUR

RESULTS

This chapter analyses the objectives of the study by giving detailed explanations on how to determine the efficacy of HIV proficiency testing panel results in a nonlaboratory-based setting in the field. It goes ahead to explain how to determine the efficacy of HIV photogrammetric testing results in a nonlaboratory-based setting in the field. In this objective, elaborate explanations are given on varied sub themes such as: validity and accuracy of photogrammetric testing, features of proficiency, reliability, repeatability and reproducibility testing as well as measures of association. The findings of the study are also vividly captured.

4.1 Determining the efficacy of HIV proficiency testing panels results in a nonlaboratory-based setting in the field

4.1.1 The validity of serological testing

The validity of serological testing was done using panel sera and the samples were tested using Determine® test kits. A 2 x 2 contingency table was created and specificity, sensitivity, negative predictive and positive predictive values were determined. The findings revealed that specificity value was low (70.1%) while the sensitivity value was high at 98.1%. Table 4.1 below shows the validity of serological testing.

Table 4.1: Validity of serological testing

	"Gold standard" (panels)	
Operators	**Positive**	**Negative**
Positive	203	8
Negative	4	19
Specificity	70.37%	
Sensitivity	98.07%	
Negative predictive	82.61%	
Positive predictive	96.21%	
Accuracy	94.87%	

4.1.2 Accuracy on proficiency testing

The overall accuracy on proficiency testing was calculated as the percentage of correct true positives and true negatives divided by the entire sample size. Therefore, (TP 203 + TN 19)/234 = 94.87%

4.1.3 Receiver Operating Characteristic and the Area Under the Curve on proficiency testing

The Receiver Operating Characteristic (ROC) curve on proficiency testing (see Figure 4.1 was constructed by plotting the true positive rate (TPR, y-axis) against the false positive rate (FPR, x-axis). The Area Under the Curve (AUC) represented a probability of 0.842 which was above the minimum threshold of 0.50 as shown in Figure 4.1: Reciever Operator Characteristic curve on proficiency testing

The coordinates of the proficiency testing are based on the law of certainty. The closer AUC is to 1, the better the model. Therefore Table 4.2 indicates that the model for proficiency testing was good.

Table 4.2, while the coordinates on the ROC curve are shown in Table 4.3

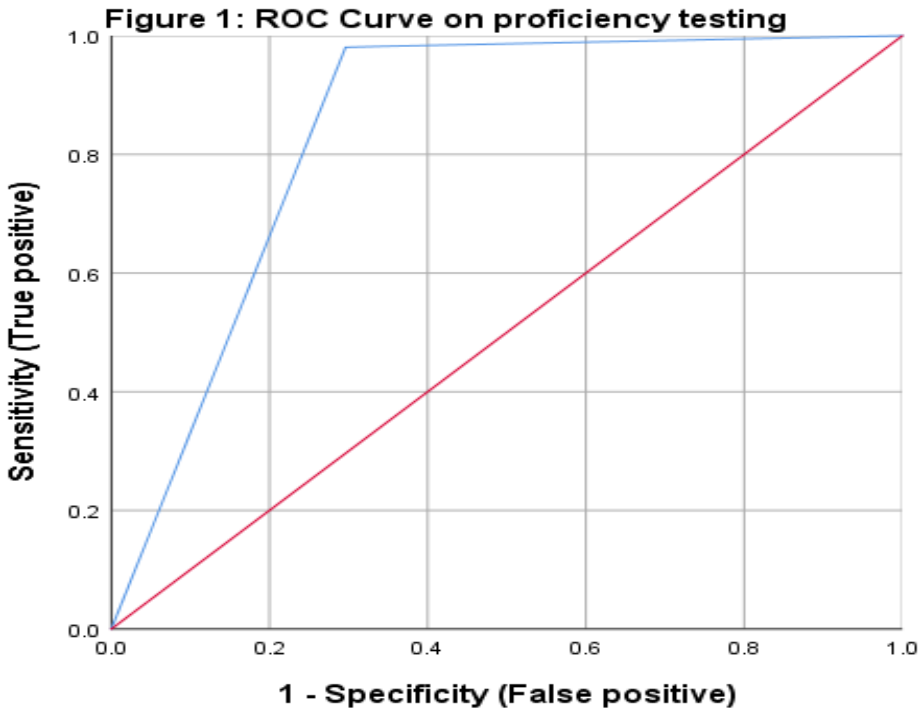

Figure 4.1: Reciever Operator Characteristic curve on proficiency testing

The coordinates of the proficiency testing are based on the law of certainty. The closer AUC is to 1, the better the model. Therefore Table 4.2 indicates that the model for proficiency testing was good.

Table 4.2: Area under the ROC curve on proficiency testing

Test Result Variable(s): panels				
Area	**SE**	**p-value**	**LB**	**UB**
0.842	0.054	< 0.001	0.737	0.947

Notes: Null hypothesis: true area = 0.5. LB =Lower Bound, UB=Upper Bound both at 95% Confidence Interval

The sensitivity value on proficiency testing as shown in the coordinate Table 4.3 was high as it represented 98.1% of the probability of predicting a positive outcome when indeed the DTS panel result was positive.

Table 4.3: Coordinates of the curve on proficiency testing

Test Result Variable(s): panels		
Positive if Greater Than or Equal To	**Sensitivity**	**1 - Specificity**
0	1	1
1.5	0.981	0.296
3	0	0

4.2 Determining the efficacy of HIV photogrammetric testing results in a nonlaboratory-based setting in the field

4.2.1 The validity of photogrammetric testing

The validity of photogrammetric testing was done and its finding represented in 2 x 2 contingency table. The specificity, sensitivity, negative predictive and positive predictive values were determined as shown in Table 4.4. The study revealed that the sensitivity rate on photogrammetric testing was higher (96.4%) than the specificity rate (70%).

Table 4.4: Validity of photogrammetric testing

	"Gold standard" (Photos)	
Operators	Positive	Negative
Positive	216	3
Negative	8	7
Specificity	70%	
Sensitivity	96.43%	
Negative predictive	46.67%	
Positive predictive	98.63%	
Accuracy	95.30%	

4.2.2 Accuracy of photogrammetric testing

The overall accuracy of photogrammetric testing was calculated as the percentage of correct true positives and true negatives divided by the entire sample size. Therefore, (TP 216 + TN 7)/234 = 95.30%

4.2.3 Receiver Operating Characteristic and the Area Under the Curve on photogrammetric testing

The Receiver Operating Characteristic (ROC) curve on photogrammetric testing (see Figure 4.2) was constructed by plotting true positive rate (TPR, y-axis) against the false positive rate (FPR, x-axis). The Area Under the Curve (AUC) represented a probability of 0.832 as shown in Table 4.5 while the coordinates on the ROC curve are shown in Table 4.6. The study noted that AUC value on photogrammetric testing was good as it represented 83.2% which was far above the minimum of 50%.

Table 4.5: Area under the ROC curve on photogrammetric testing

Test Result Variable(s): photos				
Area	**SE**	***p*-value**	**LB**	**UB**
0.832	0.088	< 0.001	0.66	1

Notes: Null hypothesis: true area = 0.5. LB =Lower Bound, UB=Upper Bound both at 95% Confidence Interval

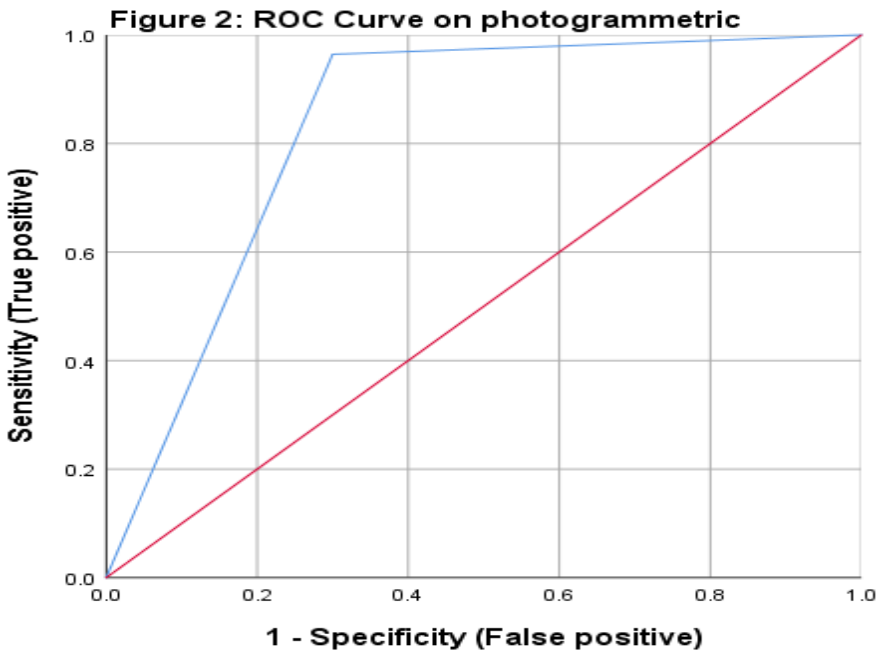

Figure 4.2: Receiver Operator Characteristic curve of photogrammetric testing

The sensitivity value on proficiency testing as shown in the coordinate Table 4.6, was high as it represented 96.4% of the probability of predicting a positive outcome when indeed the DTS panel result was positive.

Table 4.6: Coordinates of the Curve on photogrammetric testing

Test Result Variable(s): photos **Positive if Greater Than or Equal To**	**Sensitivity**	**1 – Specificity**
0	1	1
1.5	0.964	0.3
3	0	0

4.3 Determining the efficacy of HIV photogrammetric testing results in a nonlaboratory-based setting in the field

4.3.1 Sample characteristics of proficiency testing and photogrammetric testing

The sample characteristics were determined using simple descriptive statistics. The operators were categorised based on the type of their workstations, that is, the type of the HIV testing facility that they were attached to. Table 4.7 below categorises the

operators based on the type of their workstations. As captured in Table 4.7, the number of operators participating in the PT rounds increased from round 9 to round 11. This was an indication of the increase in the enrolment of individuals participating in the PT scheme. The pattern depicted was evident in all the HIV testing facilities.

Table 4.7: Operators' participation in proficiency Testing rounds

Type of HIV testing facility	Proficiency testing rounds	Frequency	Per cent
HBTC (*n* =20)	Round 9	2	10
	Round 10	6	30
	Round 11	12	60
LAB (*n* = 50)	Round 9	14	28
	Round 10	17	34
	Round 11	19	38
PITC (*n* = 53)	Round 9	15	28.3
	Round 10	18	34
	Round 11	20	37.7
PMTCT (*n* =40)	Round 9	12	30
	Round 10	13	32.5
	Round 11	15	37.5
VCT (*n* =71)	Round 9	22	31
	Round 10	22	31
	Round 11	27	38

Notes: HBTC = Home Based Testing & Counselling, LAB = Laboratory, PITC = Provider Initiated Testing & Counselling, PMTCT = Prevention of Mother-to-Child Transmission, and VCT = Voluntary Counselling & Testing

The study noted that among the respondents who tested using rapid diagnostics tests, the highest number of those who participated in the study in almost all the testing facilities registered, were those who had experience of between five to nine years (representing 33.6%)**.** This group was closely followed by operators who had 1 to 4 years of experience in testing HIV using rapid diagnostic tests (28.9%) as shown in Table 4.8. The distribution of the operators based on the type of their workstation is also represented in Table 4.8. The Voluntary Counselling and Testing (VCT) centres

had the highest number of operators in the study while the least number came from the HBTC operators as revealed in Table 4.8

Table 4.8: Operators' experience in testing using rapid diagnostic tests

Type of health facility	less than one year	1-4 years	5-9 years	Over 10 years
HBTC	2(0.9%)	5 (2.1%)	10 (4.3%)	3 (1.3%)
LAB	11 (4.7%)	11 (4.7%)	17 (7.2%)	11 (4.7%)
PITC	11 (4.7%)	16 (6.8%)	17 (7.2%)	9 (3.8%)
PMTCT	5(2.1%)	13 (5.5%)	15 (6.4%)	7 (3%)
VCT	9 (3.8%)	23 (9.8%)	20 (8.5%)	20 (8.5%)
Overall experience	38(16.2%)	68 (28.9%)	79 (33.6%)	50 (21.3%)

Notes: HBTC = Home Based Testing & Counselling, LAB = Laboratory, PITC = Provider Initiated Testing & Counselling, PMTCT = Prevention of Mother-to-Child Transmission, and VCT = Voluntary Counselling & Testing

The distribution of the operators on the basis of the type of their district of operation is represented in Table 4.9. The study revealed that Langat'a had the highest number of operators at 12.8% followed closely by Westlands 11.9%, Kasarani 10.9%, Nairobi North at 5.1%, Embakasi at 3.4% and Njiru which had the least representation at 0.9% as captured in Table 4.9. The variation in the number of operators was dependent on the time when the PT was introduced in the districts in Nairobi. There was a direct relationship between the time the PT scheme was introduced in each district and the number of operators enrolled in the study.

Table 4.9: Distribution of operators based on DASCO District in Nairobi

District	Frequency (*n*=234)	Percent
Langata	30	12.8
Westlands	28	11.9
Kasarani	25	10.6
Makadara	24	10.2
Nairobi West	24	10.2
Dagoreti	23	9.8
Nairobi East	23	9.8
Kamukunji	18	7.7
Starehe	18	7.7
Nairobi North	12	5.1
Embakasi	8	3.4
Njiru	2	0.9

The overall scores for the proficiency and photogrammetric testing 2013 were computed using frequencies and per cent scores as shown in Table 4.10 below. Photogrammetric testing had a smaller number of overall unsatisfactory scores compared to proficiency testing. Overally, satisfactory scores surpassed unsatisfactory ones. Proficiency testing had a satisfactory score of 76.5% which was lower compared to the score on photogrammetric testing which stood at 89.3%.

Table 4.10: Sample characteristics

Test (*n* = 234)	Overall score	Frequency	Per cent
Proficiency testing 2013	Unsatisfactory	55	23.5
	Satisfactory	179	76.5
Photogrammetric 2013	Unsatisfactory	25	10.7
	Satisfactory	209	89.3

4.3.2 Reliability Testing

Inter-rater reliability testing was done using Fleiss Kappa and an overall Kappa score of 0.649 which represented significant substantial agreement was noted. The findings on proficiency testing are shown in Table 4.11. On a similar note, the findings on photogrammetric testing revealed an overall score of 0.708 which represented significant substantial agreement at 95% Confidence level as shown in Table 4.12. From the findings presented, the Fleiss Kappa on photogrammetric was higher compared to the one on proficiency testing.

Table 4.11: Fleiss kappa for proficiency testing

Rating Category	Conditional Probability	Kappa	Asymptotic Standard Error	Z	p-Value	LB	UB
Negative	0.693	0.649	0.038	17.201	< 0.001	0.575	0.723
Positive	0.956	0.649	0.038	17.201	< 0.001	0.575	0.723

Notes: Overall Fleiss kappa = 0.649. LB =Lower Bound, UB =Upper Bound. Confidence level at 95%. 0.61 – 0.80 represent substantial agreement. n = 234

Table 4.12: Fleiss kappa for photogrammetric testing

Rating Category	Conditional Probability	Kappa	Asymptotic Standard Error	Z	p-Value	LB	UB
Negative	0.725	0.708	0.038	18.769	< 0.001	0.634	0.782
Positive	0.983	0.708	0.038	18.769	< 0.001	0.634	0.782

Notes: Overall Fleiss kappa = 0.708. LB =Lower Bound, UB =Upper Bound. Confidence level at 95%.0.61 – 0.80 represent substantial agreement. *n* = 234

4.3.3 Repeatability Testing

The Kruskal-Wallis test was performed to determine if the repeatability tests on the three rounds both for the PT and Photos were significant. The preliminary findings on photogrammetric 2013 and proficiency testing 2013 showed statistically significant differences among the three test rounds on the study as shown in Table

4.13. The significance interpretations were based on Bonferroni correction ($p < 0.016$) to control for inflation of type 1 error.

Table 4.13: Kruskal-Wallis on repeatability testing

Variables	Round	N	Mean Rank
photogrammetric 2013	Round 9	65	97.6
	Round 10	76	126.92
	Round 11	94	123.71
Proficiency Testing 2013	Round 9	65	83.8
	Round 10	76	145
	Round 11	93	118.58
Test Statistics			
	photogrammetric 2013	proficiency testing 2013	
Chi-Square	27.50	53.16	
df	2	2	
Asymp. P	< 0.001	< 0.001	

Notes: Kruskal Wallis Test. The Grouping Variable: PT Round

However, the Kruskal-Wallis finding did not reveal how the actual test rounds differed. This prompted further analysis using the Mann-Whitney U test whose findings are shown in Table 4.14 Significant findings were observed between Round 9 & 10 ($p < 0.001$), Round 9 & 11 ($p < 0.001$). However, a non-significant finding was observed between Round 10 & 11 ($p = 0.374$) only on photogrammetric testing.

Table 4.14: Mann-Whitney Test on repeatability testing

Groups	Statistic	Overall proficiency testing results	Overall photogrammetric testing results
Round 9 & 10	Mann-Whitney U	1178	1851
	p value (2-tailed)	< 0.001	< 0.001
Round 9 & 11	Mann-Whitney U	2124	2348
	p value (2-tailed)	< 0.001	< 0.001
Round 10 & 11	Mann-Whitney U	2736	3437
	p value (2-tailed)	< 0.001	0.374

Notes: Grouping Variable: Proficiency testing rounds 9, 10 & 11

4.3.4 Reproducibility Testing

Reproducibility was determined based on the operators' workstations. The findings were interpreted as significant based on Bonferroni correction. The findings shown in Table 4.15 indicates that only the PITC had significant outcomes at both proficiency and photogrammetric testing. From the findings, it is evident that Photogrammetric testing had the highest number of insignificant outcomes at facility level compared to proficiency testing.

Table 4.15: Kruskal-Wallis on reproducibility testing

Facility	Test statistics	Overall proficiency testing results	Overall photogrammetric testing results
HBTC	Kruskal-Wallis H	19	5.34
	df	2	2
	p value	< 0.001	0.069
LAB	Kruskal-Wallis H	49	8.043
	df	2	2
	p value	< 0.001	0.018
PITC	Kruskal-Wallis H	52	13.722
	df	2	2
	p value	< 0.001	0.001
PMTCT	Kruskal-Wallis H	7.887	5.353
	df	2	2
	p value	0.019	0.069
VCT	Kruskal-Wallis H	28.018	5.604
	df	2	2
	p value	< 0.001	0.061

Notes: Kruskal Wallis Test conducted, and the grouping variable was proficiency testing rounds. HBTC = Home-Based Testing & Counselling, LAB = Laboratory, PITC = Provider Initiated Testing & Counselling, PMTCT = Prevention of Mother-to-Child Transmission, and VCT = Voluntary Counselling & Testing. The p-value used to interpret the findings was Bonferroni corrected $p = 0.016$. That is 0.05 /3 groups. ** means significant at 0.016.

The findings on the Mann-Whitney test as shown in Table 4.16, indicate that most significant results were observed in Round 9 and Round 10 for proficiency testing. Most of the findings on photogrammetric testing were insignificant apart from the finding on HBTC, Provider Initiated Testing & Counselling (PITC). In the

Prevention of Mother-to-Child Transmission (PMTCT), no significant findings were observed in Round 10 for both proficiency and photogrammetric testing.

Table 4.16: Mann-Whitney Test on reproducibility testing

Facility	PT_9	Photos_9	PT_10	Photos_10	PT_11	Photos_11
HBTC	0.008	0.693	< 0.001	0.014	1	0.039
LAB	< 0.001	0.048	< 0.001	0.037	1	1
PITC	< 0.001	0.009	< 0.001	0.006	1	1
PMTCT	0.06	0.026	0.256	0.223	0.005	0.18
VCT	1	0.019	< 0.001	0.279	< 0.001	0.11

Notes: Kruskal Wallis Test conducted, and the grouping variable was proficiency testing rounds. HBTC = Home-Based Testing & Counselling, LAB = Laboratory, PITC = Provider Initiated Testing & Counselling, PMTCT = Prevention of Mother-to-Child Transmission, and VCT = Voluntary Counselling & Testing. The *p*-value used to interpret the findings was Bonferroni corrected $p = 0.016$. That is 0.05 /3 groups. ** means significant at 0.016. PT = Proficiency testing

4.3.5 Measures of association

A 2 x 2 contingency table and Chi-Square statistics measured the association between the operators' results and the gold standard on proficiency testing as shown in Table 4.17. There was a significant association between the findings of operators and the gold standard results on proficiency testing, X^2 (1, N = 234) = 126.22, $p <$ 0.00001 at 95% confidence level.

Table 4.17: Measurement of the association on proficiency testing

	"Gold standard" (Panels)	
Operators	Positive	Negative
Positive	203	8
Negative	4	19

A 2 x 2 contingency table and Chi-Square statistics measured the association between the operators' results and the gold standard on photogrammetric testing as shown in Table 4.18. There was a significant association between the findings of

operators and the gold standard results on photogrammetric testing, X^2 (1, N = 234) = 70.41, p < 0.00001 at 95% confidence level.

Table 4.18: Measurement of the association on photogrammetric testing

	"Gold standard" (photos)	
Operators	Positive	Negative
Positive	216	3
Negative	8	7

A 2 x 2 contingency table and Chi-Square statistics as shown in Table 4.19 indicated that there was a significant association between the findings of proficiency and those of photogrammetric testing, X^2 (1, N = 234) = 6.54, p < 0.01 at 95% confidence level.

Table 4.19: Overall Measurement of association

	Proficiency Testing		
Photogrammetric Testing	Satisfactory	Unsatisfactory	Totals
Satisfactory	165	44	209
Unsatisfactory	14	11	25
Totals	179	55	234

CHAPTER FIVE

DISCUSSION

The first objective determined the efficacy of HIV Proficiency Testing panels results in a nonlaboratory-based setting in the field. The validity of serological testing using panel sera had exceptionally good results with a higher sensitivity (98.07%) and positive predictive value (PPV) of 96.21%. On the contrary, it had a lower specificity (70.37%) and negative predictive value (NPV) of 82.61%. The sensitivity value of 98.07% was an indication that the operators were able to correctly identify positive test samples from the negative ones which had a specificity value of 70.37%. The was no significant difference in the sensitivity value between the current study and the one that appears on the package insert of Determine™ HIV-1/2 Ag/Ab combo testing kit. Moreover, the findings of this evaluation were consistent with those documented in other studies that had reported higher sensitivity values after using Determine™ HIV-1/2 Ag/Ab combo testing kit. Some of the studies whose findings corroborated the current one are as follows: Van den Berk *et al.* (2003); Lyamuya *et al.* (2009); Rosenberg *et al.* (2012); Kroidl *et al.* (2012); Granade *et al.* (2005); Koblavi-Dème *et al.* (2001) and Ménard *et al.* (2005) among others. Therefore, the findings of the current study were able to replicate the results they were intended to portray. For instance, that Determine™ HIV-1/2 Ag/Ab combo testing kit is a useful test for the detection of HIV antibodies especially in resource-limited settings. In addition, this kit can be used to test competencies of operators during proficiency testing schemes. The reason why a higher sensitivity value is preferred is because it minimises the chances of generating false negative results thus ensuring that it can detect low titre antibodies at the early stages of infection.

The diagnostic performance of a test which has a binary predictor is assessed using the measures of sensitivity and specificity. WHO (2004) guidelines in resource-limited settings recommend the use of RDT's which have a sensitivity score greater than 99%. The sensitivity performance of most commercialized HIV rapid tests as shown in their package inserts are exceptional and remarkably close to 100%. However, the sensitivity and specificity results as indicated in a package insert of a

test kit, are not necessarily achieved in practice (Boadu *et al.*, 2016). This observation corroborates the one made in the current study with regard to variation in terms of experience among operators.

Lower specificity is a sign of occurrence of high cross-reactivity when the serum is used as the test specimen. The result of the current evaluation is slightly lower compared to the studies mentioned above which showed sensitivity rates greater than 99%. Some of the possible reasons could be that in the current study the operators tested serum samples (DTS) while in those previously mentioned whole blood samples were used. It is also worth noting that there are variations in technical expertise among operators a factor that is also likely to contribute to low specificity. Moreover, proficiency testing scheme is a new program in the country therefore operators have limited know how in relation to how it is conducted a factor that may contribute to low specificity rates.

The positive predictive value for Determine™ HIV-1/2 Ag/Ab combo testing kit was 86.26% (95% confidence interval 46.96% to 97.80%), while the negative predictive value was 99.99% (95% confidence interval 99.98% to 100%). The positive predictive value for the current study was 96.21% (95% confidence interval 93.41 to 97.84%) which indicates that the results posted by the operators and the gold standard were able to match. This means that the operators had 96.21% chance of identifying the positive sample during proficiency testing. The validation of proficiency testing correctly identified most (203 out of 234) positive tests and the finding was not problematic because 8 operators identified a positive test as a negative one. The performance of any HIV rapid test kit is mainly influenced by positive and negative predictive values of the kits used. Several studies have reported positive predictive values for example, Kroidl *et al.* (2012) whole blood sample 32.91% and plasma 82.57%; Rosenberg *et al.* (2012) whole blood sample 97.0%; and Singer *et al.* (2005) serum 55.4%. The positive predictive values recorded by the above-mentioned studies that used serum were exceptionally lower as compared to the findings of the current study. The negative predictive value of 82.61% (95% confidence interval 63.58 to 92.82%) indicated that the operators were able to match the negative test as predicted by the Gold standard during proficiency

testing. The finding was very encouraging because only 4 operators out of the possible 234 identified a negative test as a positive one.

According to Cook (2000), positive and negative predictive values are very important especially when it comes to screening. This is because the positive predictive value indicates the probability that a disease is present while a negative predictive value indicates lack of the disease. However, these predictive values can change dramatically based on the background frequency of the disease. Cook (2000) states the predictive value to be a function of sensitivity and specificity of the test and the prevalence of a disease. This means that as the prevalence drops, so does the positive predictive value and vice versa.

The area under the curve (AUC) was used to summarise the performance of the operators and the gold standard. In the current study, the AUC on proficiency testing indicated a strong probability (0.843) that a randomly chosen positive test was ranked higher than a randomly chosen negative test. This suggests an 84.3% chance that the operators testing the DTS panels correctly distinguish between a positive test result and a negative one. According to Hosmer Jr *et al.* (2013), the AUC ranging between 0.80-0.90 represents excellent discrimination which represents good test accuracy. Thus, the validation of serology for the proficiency testing in the current study represented a good test accuracy with a higher predictive ability of the positive DTS compared to negative DTS. Additionally, it is important to note that the ROC curve was representing the tradeoff between sensitivity and specificity such that any increase in sensitivity was accompanied by a decrease in specificity. Thus, in the current study, sensitivity was 98.07% while specificity was 70.37%. In HIV testing, a diagnostic kit with high sensitivity matters most because it is an indication of minimum false negatives. Kroidl *et al.* (2012) reported AUC values for whole blood sample as 0.984 and plasma 0.989. These two values were higher than the one reported in the current study. The variation is mainly due to differences in the samples that were used in the screening process. While in the current study individuals were screened for HIV using serum samples, the study by Kroidl *et al.* (2012) used whole blood samples.

The second objective determined the efficacy of HIV photogrammetric testing results in a nonlaboratory-based setting in the field. The sensitivity value of photogrammetric testing was 96.43% (95% confidence interval 93.08 to 98.45%) with a positive predictive value (PPV) of 98.6%, and 70% and 46.7% for specificity and negative predictive value respectively. Although there are no similar studies to do comparisons with, it is crucial to note that the sensitivity value of the current study was slightly lower than that of Determine™ HIV-1/2 Ag/Ab combo testing kit as indicated in the package insert 99.9% (95% confidence interval 99.4 to 100%). This means that the current study was able to identify a gap in the sensitivity of photogrammetric testing and addressed it. Future studies will use the sensitivity value for the current study as their benchmark for photogrammetric testing in HIV.

The current study revealed that the specificity value for photogrammetric testing was 70% (95% confidence interval 34.75 to 93.33%) while the specificity value of Determine™ HIV-1/2 Ag/Ab combo testing kit as indicated in the package insert was 99.84% (95% confidence interval 99.11% to 100%). Although these two values cannot be directly compared, it is evident that the value of the current study was low. However, it is worth noting that the two studies were conducted using different specimens, and in varied environmental conditions. The finding for Determine™ HIV-1/2 Ag/Ab combo testing kit was achieved using excellent experimental conditions such as a good laboratory setting and qualified personnel who were evaluating the test kits. This was not the case with the Photogrammetric testing done in the current study which was conducted under a field evaluation set up. The PT scheme was also conducted in the months which were humid and rainy therefore, these unfavourable weather conditions could have compromised the quality of the photo images. It is worth pointing out that no study had addressed the specificity rate of photogrammetric testing. Thus, the current study fills the knowledge gap by providing a foundation upon which future studies on HIV rapid test kits can draw literature.

There is also a literature gap in the positive and negative predictive values on photogrammetric testing. No study had documented these values on photogrammetric testing of HIV using Determine test kits. Findings from the current

study show that there is a positive predictive value of 98.63% (95% confidence interval 96.54 to 99.46%), and a negative predictive value of 46.67 (95% confidence interval 28.38 to 65.90%). However, both positive and negative predictive values are dependent on the prevalence of the disease. Even though there is no literature documentation on AUC and ROC curves on photogrammetric testing of HIV, significant findings were revealed in the current study in relation to the AUC and the ROC curves. This means that this study was able to identify and fill the literature gaps in the two values. Future studies will now have values that they can use to benchmark when conducting research on photogrammetric testing. This information will also provide as a comprehensive guideline to be used in identifying operators who urgently require HIV testing refresher courses. Photogrammetric testing is a new concept that incorporates digital or hardcopy photos as complementary training tools that can be used to improve interpretation of test results. The technique has been used successfully to improve test results that highly depend on visual acuity especially in the field of radiology. There is limited literature and data regarding the use of photogrammetric in the training of HIV operators who use their visual acuity to interpret rapid antigen/antibody test, rapid antibody screening test, or oral fluid antibody self-test all of which depend on visual acuity.

The third objective assessed the feasibility of using photographed rapid HIV test results as a monitoring tool for testing competencies and training operators in non-laboratory-based settings. The accuracy of a HIV test kit is extremely critical for the success of the rapid expansion of HIV/AIDS detection, prevention, monitoring and treatment in resource-limited settings. The accuracy rate of Determine™ HIV-1/2 Ag/Ab combo testing kit was 99.83% (95% confidence interval of 99.48% to 99.97%) while the current study had an accuracy rate of 94.87% (95% confidence interval 91.21% to 97.32%) and a false-positive rate of 3.4%. In a study carried out by Anzala *et al.* (2008) in Kenya and Uganda, accuracy rates for the two study sites in Kenya (Kangemi and Kilifi) were both reported to be 99.60% while the other two sites in Uganda (Kakira and Masaka) had accuracy rates of 90.49% and 95.71% respectively. The accuracy rate in the current study was slightly lower compared to the ones found in Anzala *et al.* (2008) because of the differences in the samples

tested. The current study used PT DTS panels made up of serum while Anzala *et al.* (2008) used whole blood sample from volunteers.

The accuracy rate of photogrammetric testing which was determined as 95.30% (95% confidence interval 91.74% to 97.63%) was compared with the finding reported by Learmonth *et al.* (2008) which reported an accuracy rate of 80% and a false-positive interpretation of 11.5%. This means that the current study has a higher accuracy rate compared to that of Learmonth *et al.* (2008) and thus sets a higher accuracy benchmark value when it comes to determining the efficacy of HIV photogrammetric testing results in non-laboratory based settings in the field. The differences in the two studies can be attributed to the fact that the study by Learmonth *et al.* (2008) involved 191 laboratories spread all over the world with different environmental conditions that would probably affect the quality of photographs on hard copies. Moreover, the participants in the study had different levels of training, visual acuity and experiences in using RDTs. As for the current study, the higher level of accuracy on photogrammetric testing was attributed to the study being conducted in one location that is, Nairobi County as such, the environmental conditions for all the participants and participating sites were similar.

WHO (2004) guidelines in resource-limited settings recommend the use of RDT's which have a sensitivity score greater than 99%. In the current study, the sensitivity of PT was 98.07% while that of photogrammetric testing was 96.43%. One of the possible explanations for the slightly lower sensitivity rate was the short supplies of test kits which limited the number of operators participating in PT. The short supplies was due to poor forecasting which was rampant in Kenya since 2010 as documented by the National AIDS and STI Control Programme (NASCOP) & Ministry of Public Health and Sanitation (2012). Most facilities had adopted a random rotational method to allow every member in their testing facility to participate in PT schemes. This compromised not only the experience in testing for HIV using rapid test kits but also the experience in participating in PT schemes. Therefore, in terms of validity, this study reveals closer similarities in using HIV Proficiency Testing panels and photogrammetric testing in non-laboratory setting. It

further exemplifies the potential for using photographed rapid HIV test results as a monitoring tool for testing competencies and training operators in such settings.

5.1 Conclusion

This study concludes the following;

1 Determine algorithm is still sensitive and specific and it can continue being used for proficiency testing of HIV panels especially in limited-resource settings where operators testing competencies is vital. However, it is also noted that such usage should be accompanied by continuous quarterly participation in PT schemes.

2 The higher accuracy rate of photogrammetric testing sets a benchmark value when it comes to determining the efficacy of HIV photogrammetric testing results in a nonlaboratory-based setting in the field. The accuracy of interpretation of HIV test results can be used to assess proficiency levels of test operators. Nevertheless, photogrammetric tests are subjectively read and interpreted as such, they are open to variations.

3 Photogrammetric testing can be used as a complementary training tool to identify operators who urgently require HIV testing refresher courses. This will cut down on the entire cost of refresher training programs since only operators with low photogrammetric scores will be recommended for further HTC training programs. However, it is important to note that photogrammetric testing is not a total replacement for proficiency testing but a complementary tool that is intended to improve on the outcomes of nonlaboratory based testing where RDT's are used in resource-limited settings.

5.2 Recommendations

The study recommends the following;

1 Determine test kit used in proficiency testing of HIV is still effective in testing operators competencies in limited-resource settings.

2 Photogrammetric testing incorporated in routine PT schemes can reduce the variations witnessed among operators and improve the accuracy of interpretations of Determine test kits thereby boosting the credibility of the HIV testing process.

3 The future for photogrammetric testing lies in the full adoption of ICT services which include the integration of mobile applications and websites into the PT schemes. In general, the National government needs to link all HIV testing facilities with affordable internet services. This will enhance the adoption of appropriate solutions that will integrate photogrammetric testing into routine proficiency testing in Kenya. The NHRL can be supported to digitise their systems in readiness for the incorporation of photogrammetric testing as part of their routine training tool in relation to proficiency testing. The outcomes of this study optimistically provide a starting point for integrating photogrammetric testing to the routine proficiency testing not only within Nairobi County but also nationally.

REFERENCES

Abokyi, L., Mahama, C., Sulemana, A., Adda, R., Amenga-Etego, S., Baiden, F., & Owusu-Agyei, S. (2014). Willingness to undergo HIV testing in the Kintampo Districts of Ghana. *Ghana Medical Journal*, *48*(1), 43-46.

Altman, D. G., & Bland, J. M. (2010). Statistical notes *British Medical Journal*.

Anzala, O., Sanders, J., Kamali, A., Katende, M., Mutua, G., Ruzagira, E., Stevens, G., Simek, M., & Price, M. (2008). Sensitivity and specificity of HIV rapid tests used for research and voluntary counselling and testing. *East African Medical Journal*, *85*(10), 500-504.

Black, V., Von Mollendorf, C., Moyes, J., Scott, L., Puren, A., & Stevens, W. (2009). Poor sensitivity of field rapid HIV testing: implications for mother-to-child transmission programme. *BJOG: An International Journal of Obstetrics and Gynaecology*, *116*(13), 1805-1808.

Boadu, R., Darko, G., Nortey, P., Akweongo, P., & Sarfo, B. (2016). Assessing the sensitivity and specificity of First Response HIV-1-2 test kit with whole blood and serum samples: a cross-sectional study. *AIDS Research and Therapy*, *13*(1), 9.

Boelaert, M., Hattacharya, S. B., Chappuis, F., Safi , S. H. E., Haiilu, A., Mondal, D., Rijab, S., Sundar, S., Wasunna, M., & Peeling, R. W. (2007). Evaluation of rapid diagnostic tests. Visceral leishmaniasis. *Nature Publishing Group*, *5*(11), S30-S39.

Bott, S., Neuman, M., Helleringer, S., Desclaux, A., Asmar, K. E., Obermeyer, C. M., & Group, M. S. (2015). Rewards and challenges of providing HIV testing and counselling services: health worker perspectives from Burkina Faso, Kenya and Uganda. *Health Policy and Planning*, *30*(8), 964-975.

Buttò, S., Suligoi, B., Fanales-Belasio, E., & Raimondo, M. (2010). Laboratory diagnostics for HIV infection. *Annali dell'Istituto Superiore di Sanità*, *46*, 24-33.

Carrio, A., Sampedro, C., Sanchez-Lopez, J. L., Pimienta, M., & Campoy, P. (2015). Automated low-cost smartphone-based lateral flow saliva test reader for drugs-of-abuse detection. *Sensors*, *15*(11), 29569-29593.

Chakrabarty, P., Rudra, S., Hossain, M., Begum, S., Mirza, T., & Rudra, M. (2015). Rapid Immuno-Chromatographic Assay for the Detection of Antibodies to HIV Compare with Elisa among Voluntary and Replacement Blood Donor of Mymensingh Medical College Hospital. *Mymensingh Medical Journal*, *24*(2), 276.

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., Ferrero, E., Agapow, P.-M., Zietz, M., & Hoffman, M. M. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface*, *15*(141), 20170387.

Cook, R. R. (2000). The importance of test validity and predictive values to screening programs. *Jurimetrics*, *41*(1), 111-120. http://www.jstor.org/stable/29762690

De Fauw, J., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., Askham, H., Glorot, X., O'Donoghue, B., & Visentin, D. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, *24*(9), 1342-1350.

Doan, M., & Carpenter, A. E. (2019). Leveraging machine vision in cell-based diagnostics to do more with less. *Nature materials*, *18*(5), 414-418.

Dubey, A. (2016). Applications of machine learning: cutting edge technology in HIV diagnosis, treatment and further research. *Computational Molecular Biology*, *6*.

Eller, L. A., Eller, M. A., Ouma, B. J., Kataaha, P., Bagaya, B. S., Olemukan, R. L., Erima, S., Kawala, L., de Souza, M. S., & Kibuuka, H. (2007). Large-scale human immunodeficiency virus rapid test evaluation in a low-prevalence ugandan blood bank population. *Journal of Clinical Microbiology*, *45*(10), 3281-3285.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118.

Gotch, F., Holmes, H., & Imami, N. (2005). The importance of standardization of laboratory evaluations in HIV vaccine trials. *Microbes and Infection 7*, 1424-1432.

Granade, T. C., Parekh, B. S., Tih, P. M., Welty, T., Welty, E., Bulterys, M., Ndikintum, G., Nkuoh, G., & Tancho, S. (2005). Evaluation of rapid prenatal human immunodeficiency virus testing in rural Cameroon. *Clinical and Diagnostic Laboratory Immunology*, *12*(7), 855-860.

Gray, R. H., Makumbi, F., Serwadda, D., Lutalo, T., Nalugoda, F., Opendi, P., Kigozi, G., Reynolds, S. J., Sewankambo, N. K., & Wawer, M. J. (2007). Limitations of rapid HIV-1 tests during screening for trials in Uganda: diagnostic test accuracy study. *British Medical Journal*, *335*(7612), 188.

Hamilton, C., Okoko, D., Tolhurst, R., Kilonzo, N., Thaobalt, S., & Taegtmeyer, M. (2008). Potential for abuse in the V.C.T Counseling room: service provider's perceptions in Kenya. London school of hygiene and tropical medicine. *Health policy and planning 23*, 390-396.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.

Jaspard, M., Le Moal, G., Saberan-Roncato, M., Plainchamp, D., Langlois, A., Camps, P., Guigon, A., Hocqueloux, L., & Prazuck, T. (2014). Finger-stick

whole blood HIV-1/-2 home-use tests are more sensitive than oral fluid-based in-home HIV tests. *PloS One, 9*(6), e101148.

Kate, M. L., Dale, A. M., Darren, K. J., Sandy, K. W., Thein-Thein, A., & Elizabeth, M. D. (2008). Assessing proficiency of interpretation of Rapid Immunodificiency Virus Assays in Non-laboratory settings:Ensuring Quality of Testing. *Journal of Clinical Microbiology, 46*(5), 1692-1697.

Koblavi-Dème, S., Maurice, C., Yavo, D., Sibailly, T. S., Kabran, N., Kamelan-Tano, Y., Wiktor, S. Z., Roels, T. H., Chorba, T., & Nkengasong, J. N. (2001). Sensitivity and specificity of human immunodeficiency virus rapid serologic assays and testing algorithms in an antenatal clinic in Abidjan, Ivory Coast. *Journal of Clinical Microbiology*, *39*(5), 1808-1812.

Kroidl, I., Clowes, P., Mwalongo, W., Maganga, L., Maboko, L., Kroidl, A. L., Geldmacher, C., Machibya, H., Hoelscher, M., & Saathoff, E. (2012). Low specificity of determine HIV1/2 RDT using whole blood in south west Tanzania. *PloS One, 7*(6), e39529.

Lab tests online. (2011, March 22, 2011). *How reliable is laboratory testing?* . American Association of clinical chemistry. Retrieved 10th July from http://labtestsonline.org/understanding/features/reliability.html

Learmonth, K. M., McPhee, D. A., Jardine, D. K., Walker, S. K., Aye, T.-T., & Dax, E. M. (2008). Assessing proficiency of interpretation of rapid human immunodeficiency virus assays in nonlaboratory settings: ensuring quality of testing. *Journal of Clinical Microbiology*, *46*(5), 1692-1697.

lieu, P., Yang, H.-T., Qiang, L. Y., JIN, H., Xiao, S., & Shi, Z.-x. (2010). Evaluation of 30 commercial assays for the detection of antibodies to HIV in China using classical and Bayesian statistics. *Journal of virological methods 170*, 73-79.

Lyamuya, E. F., Aboud, S., Urassa, W. K., Sufi, J., Mbwana, J., Ndugulile, F., & Massambu, C. (2009). Evaluation of simple rapid HIV assays and

development of national rapid HIV test algorithms in Dar es Salaam, Tanzania. *BMC Infectious Diseases*, *9*(1), 19.

Makuwa, M., Souquière, S., Niangui, M.-T., Rouquet, P., Apetrei, C., Roques, P., & Simon, F. (2002). Reliability of rapid diagnostic tests for HIV variant infection. *Journal of Virological Methods*, *103*(2), 183-190.

Mandrekar, J. N. (2010a). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, *5*(9), 1315-1316.

Mandrekar, J. N. (2010b). Simple statistical measures for diagnostic accuracy assessment. *Journal of Thoracic Oncology*, *5*(6), 763-764.

Mayhood, M. K., Afwamba, I. A., Odhiambo, C. O., Ndanu, E., Thielman, N. M., Morrissey, A. B., Shao, J. F., Pence, B. W., & Crump, J. A. (2008). Validation, performance under field conditions, and cost-effectiveness of Capillus HIV-1/HIV-2 and determine HIV-1/2 rapid human immunodeficiency virus antibody assays using sequential and parallel testing algorithms in Tanzania. *Journal of Clinical Microbiology*, *46*(12), 3946-3951.

Ménard, D., Maïro, A., Mandeng, M.-J., Doyemet, P., Koyazegbe, T. d. A., Rochigneux, C., & Talarmin, A. (2005). Evaluation of rapid HIV testing strategies in under equipped laboratories in the Central African Republic. *Journal of Virological Methods*, *126*(1-2), 75-80.

Moodley, D., Moodley, P., Ndabandaba, T., & Esterhuizen, T. (2008). Reliability of HIV rapid tests is user dependent. *SAMJ: South African Medical Journal*, *98*(9), 707-709.

Muchiri, S. M., Osman, S., Night, D., Muia, E., Kaiguri, P., Kimotho, J., & Ochwoto, M. (2016). Performance of selected HIV testing centers in a HIV Proficiency Testing Scheme in Kenya: a case study. *African journal of pharmacology and therapeutics*, *5*(1).

Mwangi, J., Joyceline, K., Nancy, L., Muriuki, J., Emma, M., Angella, S., Peter, M., Janet, O., & Carol, M. (2012). *Quality Management systems in HIV and AIDS: A Handbook for HTC Programs in Developing Countries*.

NACC. (2021). *Kenya World AIDS Day: Progress Report 2013-2021* (End Teenage Pregnancy Towards eliminating inequalities and new HIV infections Issue. Retrieved fom https://nacc.or.ke/wp-content/uploads/2022/02/WAD2021Report.pdf

Parekh, B. S., Anyanwu, J., Patel, H., Downer, M., Kalou, M., Gichimu, C., Keipkerich, B. S., Clement, N., Omondi, M., & Mayer, O. (2010). Dried tube specimens: a simple and cost-effective method for preparation of HIV proficiency testing panels and quality control materials for use in resource-limited settings. *Journal of Virological Methods, 163*(2), 295-300.

Piwowar-Manning, E. M., Tustin, N. B., Sikateyo, P., Kamwendo, D., Chipungu, C., Maharaj, R., Mushanyu, J., Richardson, B. A., Hillier, S., & Brooks Jackson, J. (2010). Validation of rapid HIV antibody tests in 5 African countries. *Journal of the International Association of Physicians in AIDS Care*, *9*(3), 170-172.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, *380*(14), 1347-1358.

Rosenberg, N. E., Kamanga, G., Phiri, S., Nsona, D., Pettifor, A., Rutstein, S. E., Kamwendo, D., Hoffman, I. F., Keating, M., & Brown, L. B. (2012). Detection of acute HIV infection: a field evaluation of the determine® HIV-1/2 Ag/Ab combo test. *Journal of Infectious Diseases*, *205*(4), 528-534.

Rouet, F., Ekouevi, D. K., Inwoley, A., Chaix, M.-L., Burgard, M., Bequet, L., Viho, I., Leroy, V., Simon, F., & Dabis, F. (2004). Field evaluation of a rapid human immunodeficiency virus (HIV) serial serologic testing algorithm for diagnosis and differentiation of HIV type 1 (HIV-1), HIV-2, and dual HIV-1-

HIV-2 infections in West African pregnant women. *Journal of Clinical Microbiology*, *42*(9), 4147-4153.

Singer, D. E., Kiwanuka, N., Serwadda, D., Nalugoda, F., Hird, L., Bulken-Hoover, J., Kigozi, G., Malia, J. A., Calero, E. K., & Sateren, W. (2005). Use of stored serum from Uganda for development and evaluation of a human immunodeficiency virus type 1 testing algorithm involving multiple rapid immunoassays. *Journal of Clinical Microbiology, 43*(10), 5312-5315.

Steinau, M., Swan, D. C., & Unger, E. R. (2008). Type-specific reproducibility of the Roche linear array HPV genotyping test. *Journal of clinical Virology 42*, 412-414.

Tegbaru, B., Messele, T., Wolday, D., Meles, P., Tesema, D., Birhanu, H., Tesfaye, G., Bond, K. B., Martin, R., & Rayfield, M. A. (2004). Evaluation of rapid HIV test kits on whole blood and development of rapid testing algorithm for voluntary testing and counseling centers in Ethiopia. *Ethiopian Medical Journal*, *42*(4), 267-276.

Trochim, W. M. K. (2006, 10/20/2006). *Theory of reliability*. Retrieved from http://www.socialresearchmethods.net/kb/reltypes.php

Tukei et al. (2004, September9th-12th). HIV/AIDS National research consultative meeting, Kunste Hotel-Nakuru.

Turbé, V., Herbst, C., Mngomezulu, T., Meshkinfamfard, S., Dlamini, N., Mhlongo, T., Smit, T., Cherepanova, V., Shimada, K., & Budd, J. (2021). Deep learning of HIV field-based rapid tests. *Nature Medicine*, *27*(7), 1165-1170.

Urassa, W., Nozohoor, S., Jaffer, S., Karama, K., Mhalu, F., & Biberfeld, G. (2002). Evaluation of an alternative confirmatory strategy for the diagnosis of HIV infection in Dar Es Salaam, Tanzania, based on simple rapid assays. *Journal of Virological Methods*, *100*(1-2), 115-120.

Van den Berk, G., Frissen, P., Regez, R., & Rietra, P. (2003). Evaluation of the rapid immunoassay determine HIV 1/2 for detection of antibodies to human immunodeficiency virus types 1 and 2. *Journal of Clinical Microbiology*, *41*(8), 3868-3869.

Wang, J.-S., Kee, M.-K., Choi, B.-S., Kim, C.-W., & Kim, S. S. (2011). Evaluation of external quality assessment results for HIV testing laboratories in Korea using current analytical methods. *Clinica Chimica Acta*, *412*(11-12), 1127-1132.

WHO. (2004). Rapid HIV tests: guidelines for use in HIV testing and counselling services in resource-constrained settings.

WHO. (2005). *Guidelines for assuaring the accuracy and reliability of HIV Rapid testing. Applying a quality system approach*.

WHO. (2009). Priority interventions: HIV/AIDS prevention, Treatment, and care in the health sector.

Xu, Y., Hosny, A., Zeleznik, R., Parmar, C., Coroller, T., Franco, I., Mak, R. H., & Aerts, H. J. (2019). Deep Learning Predicts Lung Cancer Treatment Response from Serial Medical ImagingLongitudinal Deep Learning to Track Treatment Response. *Clinical Cancer Research*, *25*(11), 3266-3275.

Yao, K., Wafula, W., Bile, E. C., Chiegnsong, R., Howard, S., Demby, & Nkengasong. (2010). Ensuaring the quality of HIV rapid testing in resource poor countries using a systemic approach to training *American Journal of Clinical Pathology*, *134*(4), 568-572.

Zeh, C., Amornkul, P. N., Inzaule, S., Ondoa, P., Oyaro, B., Mwaengo, D. M., Vandenhoudt, H., Gichangi, A., Williamson, J., & Thomas, T. (2011). Population-based biochemistry, immunologic and hematological reference values for adolescents and young adults in a rural population in Western Kenya. *PloS One*, *6*(6), e21040.

Zeng, N., Wang, Z., Zhang, H., Liu, W., & Alsaadi, F. E. (2016). Deep belief networks for quantitative analysis of a gold immunochromatographic strip. *Cognitive Computation*, *8*(4), 684-692.

Zhou, X.-H., McClish, D. K., & Obuchowski, N. A. (2009). *Statistical methods in diagnostic medicine* (Vol. 569). John Wiley & Sons.

APPENDICES

Appendix I: Questionnaire

Dear colleague(s),

The researchers are engaged in a research project entitled "**Comparison of efficacy of serological and photogrammetric testing algorithms for HIV screening in nonlaboratory settings"** within Nairobi County. The project will commence in May 2014 and end in July 2014. The information gathered will go a long way in assisting the government (NHRL and NASCOP) and other stakeholders in making an important decision regarding HTC services in Nairobi County. Those who have never participated in any proficiency testing scheme will have a rare opportunity to participate in one and gauge their abilities in the diagnosis of HIV using rapid tests.

Attached is a designed questionnaire that requires you to give independent views to the questions posed therein. Please spend some time and respond to the questionnaire as accurately as possible and conduct the tests just like you would to any other routine test conducted daily. The results shall be submitted strictly within two weeks upon receiving the panel sera. The project is a longitudinal study and hence you will be supplied with the testing samples once after every three months. Conduct the test using the Determine algorithm at your disposal and fill in the results. The results for the coloured photographed HIV test should be interpreted independently like normal test results. The photos will be accompanying the PT panels quarterly.

Your safety should always come first. Treat the samples with the safety precaution that they deserve. For any assistance or enquiries please contact the principal investigator through 0725-812728 or email ireri76@gmail.com. Submission of the results via e-mail is highly encouraged to increase and gauge the use of ICT services in the study. The results form and the answers to the questionnaires should be submitted via e-mail should be typed in BOLD and highlighted in RED.

Your contribution to this project will be highly appreciated. Information provided will be treated with confidentiality.

THANK YOU.

As a Nairobian you are a stakeholder in offering HIV testing services in Nairobi County and we would like to ask you a few questions regarding your views on the state of HIV Proficiency Testing and possible solutions. These will be important as the NHRL (National HIV Reference Laboratory) plans to scale up HIV Proficiency Testing services in Nairobi County.

Kindly answer by circling your preferred responses or write where space is provided. If responding to the questionnaire via email, kindly highlight your answer in red.

Personal information: (Tick appropriately) RESPONDENT NO

1. Age Under 21 21-30 31-40 41-50 51-60 over 60
2. Sex Male Female
3. Which HIV constituency unit does your HTC site belong to?

Kamukunji Starehe Kasarani Westlands Dagoretti Langata Embakasi Njiru Makadara

4. HIV Testing and Counseling (HTC) is provided in a variety of settings. What type of setting is your facility?

   Community-based (Client-Initiated HTC) Health Facility (provider-initiated HTC)

5. What is the status of the program at your testing facility?

   VCT PITC PMTCT LAB VMMC

6. What is your designation at the testing facility?

   KRCHN KRN KRCN Lab Tech Lab Techno HTC Counselor KECN

7. How long have you worked at your present workstation?

Less than 5 years 5-10 yrs 11-15 yrs 16-20 yrs 21-25 yrs 26-30 yrs

8. Experience in performing Rapid HIV Testings.

Less than 5 years 5-10 yrs 11-15 yrs 16-20 yrs 21-25 yrs 26-30 yrs

9. How do you access internet services?

Smartphone Cybercafe Internet Network in the HTC facility

SECTION A: Validation of Test Kits: (Tick appropriately)

1. Do you know what field validation of HIV test kits is? Yes No

2. Has any field validation of HIV test kits been conducted at your site? Yes No

Do not Know

3. Do you know what post-market surveillance of HIV test kits is? Yes No

4. Has your site participated in any post-market surveillance of HIV test kits?

Yes No Don't Know

5. How often has the periodic random assessment of HIV Test Kits been conducted at your current HTC Facility?

All the times Most of the time Sometimes Rarely Never

SECTION B: On-Site Supervision visits & Mentorships. (Tick appropriately).

6. How many times in a year do the personnel from the National HIV Reference Lab/MOH visit your HIV testing site?

None Once Thrice Every three months More than 3months

7. How often do you interact with your supervisor at your HCT facility?

All the times Most of the time Sometimes Rarely Never

8. Do qualified and accredited laboratory personnel participate in the mentorship of new HTC service providers in your facility?

All the times Most of the time Sometimes Rarely Never

9. Do qualified and accredited laboratory personnel participate in the mentorship of existing HTC service providers in your facility?

All the times Most of the time Sometimes Rarely Never

SECTION C: Proficiency Testing. (TICK appropriately)

10. Do you know what HIV Proficiency Testing is? Yes No

11. Have you ever participated in any HIV Proficiency Test? Yes No

12. Has your current HTC site ever participated in any HIV Proficiency Testing?

Yes No Don't Know

13. How many times did you participate in HIV Proficiency Testing in 2012?

None One Two Three Four

14. How many times did you participate in HIV Proficiency Testing in 2013?

None One Two Three Four

15. By May 2014, how many times had you participated in HIV Proficiency Testing at your current HTC site? None One Two

16. If you never participated in HIV Proficiency Testing during 2012, 2013 periods and now in 2014, briefly state why

_____

_____

17. Would you like to participate in an HIV Proficiency Testing program in 2014?

Yes No

18. If your answer to question 17 is No, briefly give your reason

_____

_____

SECTION D: Education / Professional Training: (Tick appropriately)

19. What is your highest level of education?

None /Informal Primary Secondary Secondary + Counselling Tertiary

(Medical College) University Postgraduate)

20. Do you have any formal laboratory training? Yes No

21. Have you ever attended any formal HTC training? Yes No

22. Have you ever attended any refresher HTC training? Yes No Do not Recall

23. Can you recall attending any refresher HTC training program in 2010?

Yes No Not working in an HTC site at the time Do not recall

24. Can you recall attending any refresher HTC training program in 2011?

Yes No Not working in an HTC site at the time Do not recall

25. Can you you recall attending any refresher HTC training program in 2012?

Yes No Not working in an HTC site at the time Do not recall

26. Can you recall attending any refresher HTC training program in 2013?

Yes No Not working in an HTC site at the time Do not recall

27. Have you attended any refresher HTC training program in 2014?

Yes No Not working in an HTC site at the time Do not recall

28. If you have never attended any HTC refresher training program, briefly state your reason

_____

_____

SECTION E: Documentation of Test Kits/ Testing Procedures: (Tick appropriately).

29. How often do you document the expiry date of the test kit on the test sheet?

All the times Most of the time Sometimes Rarely Never

30. How often do you document the lot Number of test kit on the test sheet?

All the times Most of the time Sometimes Rarely Never

31. How often do you document the date of testing on the test sheet?

All the times Most of the time Sometimes Rarely Never

32. How often do you document your Operator/Analyst ID on test worksheet?

All the times Most of the time Sometimes Rarely Never

33. How often do you conduct test performance on a new batch of Rapid test kits in your facility?

All the times Most of the time Sometimes Rarely Never

34. What method do you use to enter the test results in your HTC facility?

Desktop Computer/Laptop Test sheet Workbook

SECTION F: Supply Chain Management: (Tick appropriately)

35. How often do you experience stock-outs of HTC commodities at your facility?

All the times Most of the time Sometimes Rarely Never

36. How often do you receive Rapid HIV test kits from the MOH on time?

All the times Most of the time Sometimes Rarely Never

SECTION G: Biosafety Measures: (Tick appropriately)

37. How often do you use gloves?

All the times Most of the time Sometimes Rarely Never

38. How often do you wear a laboratory coat while analyzing the test samples?

All the times Most of the time Sometimes Rarely Never

39. Have you ever undergone any laboratory biosafety measures training concerning an HTC facility since 2010?

Yes No

40. Have you ever encountered any violent reaction from a client after knowing that he/she is HIV positive?

Yes No

41. Do you have Accident and incident logbooks?

Yes No

42. Briefly state how you manage waste disposal in your HTC site.

_____

_____

43. Have you been vaccinated against Hepatitis B? Yes No

44. If Yes, which year were you vaccinated?

45. If your answer to question 43 is Yes, when was the last date that you received the booster dose?

SECTION H: Availability of documented instructions on HTC (SOP's)

Standard Operating Procedures Onsite:

46. On use, donning and removal of gloves Yes No

47. Ban on open footwear Yes No

48. Hand washing procedures Yes No

49. How to clean a spill Yes No

50. How to disinfect HIV Testing areas Yes No

51. Proper disposal of test kits Yes No

52. Signage-No unauthorized persons in HIV Testing area Yes No

53. Procedures to follow in case of accidental exposure to material that is biohazardous Yes No

SECTION I: Registration of the HTC facility: (Tick appropriately)

54. Have you ever heard of the NHRL (National HIV Reference Laboratory)?

Yes No

55. Is your HTC site registered by the NHRL (National HIV Reference Laboratory)?

Yes No Not aware

56. If your response to registration is YES would you show evidence of registration?

Registration Number Available No Registration Evidence N/A

Is there any other thing that you would like to tell the researcher concerning improving HIV testing services and HIV Proficiency Testing in Nairobi County?.....
.....
.....