UNIVERSITY OF NAIROBI
SCHOOL OF MATHEMATICS

A research project submitted to the school of mathematics in partial fulfillment of the requirements for the award of the degree of Master of Science in Biometry of the University of Nairobi.

2016.

BENJAMIN OMONDI MBOYA

MSc BIOMETRY

I56/74373/2014

Modeling rape victimization in 4 Nairobi slums using discriminant analysis

Declaration

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

…………………………………….   …………………………………

Signature   Date

This project has been submitted for examination with my approval as University supervisor.

…………………………………….   …………………………………

Signature   Date

Abbreviations:

SOC – Standard of Care,

NMNW – No Means No Worldwide,

GBV – Gender Based Violence,

BBM – Ballot Box Method,

LDA- Linear Discriminant Analysis,

GLM- General Linear Model

AIDS -Acquired Immunodeficiency Syndrome

LMIC- In low and middle-income countries

MPR- Multiple Perpetrator Rape

PCA-Principal Component Analysis

ICD- Independent Component Discriminant

DA- Discriminant Analysis

NACOSTI -The National Commission for Science, Technology and Innovation

APER- Apparent error rate

MER- Misclassification Error

MIDSA- Multidimensional Inventory of Development, Sex, and Aggression

List of Figures………………………………………………………………**Page**

List of Tables……………………………………………………Page

ABSTRACT

Purpose: This study aimed at determining whether a girl's Age, Slum of residency and Grade in school can be used as predictors of rape and also for classification of girls into different rape risk groups.

Methods: Participants were a prospective cohort of 2383 adolescent girls 13-20 years old, attending one of 35 secondary schools selected by convenience sampling from 4 informal settlements of Nairobi, that is Huruma/Mathare, Dandora, Kibera, and Mukuru. These areas were selected because of their high rates of crime and nonparticipation in the previous studies conducted by Ujamaa Africa. Fisher's linear discriminant analysis (LDA) is a popular data-analytic tool for studying the relationship between a set of predictors and a categorical response. LDA was used to analyze anonymously collected baseline data from the girls on 1) Incidences of rape 2) Slum the girls come from, 3) Ages of the girls and 4) The grade in school the girls are in. General linear model was also used to analyze the same variables to determine which of these contributes most to being sexually assaulted.

Results:

LDA prediction model created categorizes respondents into raped or not raped categories with 76.62% accuracy using Age, Slum and Grade as predictors of rape. The slum from where a girl comes from and the girl's age contributes most to the likelihood of getting raped (age, slum $p > 0.001$). 70% of girls are sexually assaulted by individuals that they know or are close to.

Conclusion:

LDA can use a combination of information on age, grade and the slum from where a girl comes from to categorize a girl into either the sexually assaulted or not groups with a 76.62% accuracy rate. The biggest contributors to the likelihood of being raped for girls are the slum from where a girl comes from, the girl's age and the grade the girl is in at school. With respect to grade, girls in form 2 are the most targeted. Girls are mostly sexually assaulted by individuals close to them as opposed to strangers. Generally, sexual desirability, Empowerment/education and Economic status of a girl or woman can be used as predictors of sexual assault with respect to the likelihood of girls being raped or sexually assaulted

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor Dr Nelson Owuor for the useful comments, remarks and engagement through the learning process of this master thesis. Furthermore I would like to thank Nixon Langat for introducing me to the topic as well for the support on the way. Also, I like to thank the participants in my survey, who have willingly shared their precious time during the process of interviewing. I would like to thank my loved ones, who have supported me throughout entire process, both by keeping me harmonious and helping me putting pieces together. I will be grateful forever for your love.

Contents

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Adolescents living in Nairobi, Kenya face high rates of both sexual assault, at between 12-25% annually[1,2,3] sexual assault is responsible for a wide variety of negative health and social outcomes, including school drop-out, physical injury or disability, and psychological trauma [3,4]. While sexual assault sometimes leads to unplanned pregnancy, these outcomes also share many risk factors, including low status of women, poverty, drug and/or alcohol abuse [5]. In Many African the countries issue of Gender Based Violence (GBV) is huge: between 16% and 59% of girls and women report being assaulted sexually at least once in their lifetime, a lot of these cases many of these occur in their childhood years. In Kenya specifically, between 11% to 46% of women report sexual assault in childhood, these numbers however vary by source.

Gender-based violence especially of the sexual kind is a serious vice that transcends racial, social, economic, and regional divides. This problem negatively affects the health, development and growth of many women, girls and boys across the globe. Over the last few years there has been growing awareness of the levels of gender based violence across the globe [2, 3]. This is not limited to the adverse consequences usually goes with it. There is therefore a dire need to design and come up with more effective and better targeted interventions.

GBV and more so violence of the sexual kind usually lead to serious health issues to victims or survivors of sexual assault. Besides the physical and emotional trauma that usually follows the attacks, heightened levels of psychological and physical medical problems have in the past been demonstrated among survivors [6]. Even though sexually transmitted infections have constantly been of great concern, risk of exposure and transmission during forced sex is more increased in parts of the globe such as Africa where there is high HIV prevalence. Kenya is an example of such a country. Sexual assault not only exposes victims to a host of sexually transmitted diseases, such as Syphilis, HIV (AIDS) and gonorrhea, but could also result in impregnation by their perpetrators [7]. These are usually unwanted adolescent pregnancies.

Adolescent girls pregnancy which are more often than not unintended is a global problem that has affected many countries over the last four decades. In low and middle-income (LMIC) countries, 19% of women get pregnant before their 18th birthday, and globally 95% of births to adolescents' girls happen in these low income countries [8]. In Kenya for instance, 36.6% of adolescent between the ages 13-19 begin childbearing, while 40-50% of those pregnancies are unintended. In Kenya alone, over 13,000 girls drop out of schools due to unintended pregnancies every year [9]. Regardless of the intentions adolescent pregnancies usually have negative health outcomes for the young mothers and the infants, which is not limited to increased risk of maternal and infant deaths. Girls who suffer this usually have lifelong missed educational opportunities and reduced earning capability [10]. Preventing adolescent pregnancy, and the related dropout, is therefore essential to improving the status and opportunities for girls in Africa and the rest of the globe.

There have been studies that have tried to look into the reasons why rape is prevalent in many countries in Africa. Despite all the research on rape causation, some reasons still remain unclear around how different factors interplay to result in victimization. One of the important areas that remain unclear is the relationship between social marginalization and rape perpetration [7]. For instance, studies have so far demonstrated that in some areas men who have taken part in multiple perpetrator rape (MPR) are usually to some extent more socially marginalized than other men and are usually relatively poorer [11]. However other studies have shown that men who take part in MPR are well educated and socially privileged [7, 11].There is substantial evidence that men who experienced trauma and other forms of abuse in their childhood are more likely to perpetrate rape. Further to this, rape perpetration and victimization is usually associated with a host of antisocial behavior [12]. Studies have demonstrated that individuals who perpetrate rape are highly likely to have been at one point members of a gang, have handled weapons, have at least once fought with other men and have or are using drugs and also abuse /abused alcohol [13].

A lot of focus has been around what causes individuals to perpetrate rape as opposed to what leads to victimization of a girl or woman. One of the most widely assumed explanation for the cause of rape reasons is not limited to mental illness irresistible impulse, drugs and alcohol abuse, but sometimes also attributed to victim precipitation [12]. With respect to irresistible impulse, people believe that men will rape because they have uncontrollable urge for immediate sexual gratification that usually follows long term sexual deprivation [15]. The problem with this assumption is that it gives the perpetrator legal argument placing the act of sexual assault as an event beyond their control, which is not true. When this legal responsibility is removed from the perpetrator, it encourages them to

go ahead and commit the offense. The victim is left with very little legal argument to get out of this. Consumption pornographic material has also been widely attributed as a cause of sexual assault [12]. A lot of blame around rape has been attributed to the victim with all these rape myths. Does the victim play a role in sexual assault? What aspects of the victims appeal to the perpetrator? Studies have tried to explain the ideal victim for perpetrators.

Other theories attempting to explain rape include 'routine activity' theory which was postulated by Felson and Cohen (1979), in their theory they argue that sexual assault happens when a motivated perpetrator, a suitable victim of sexual assault, lack of a protective safe space for the victims come into play. The protective environment in their theory is defined as presence of a protective guardian or parent. According to their theory, rape or sexual assault will happen when the motivated criminal finds a victim in a situation where they are not protected [3, 14]. The reverse is also true.

This theory also assumes that males are the perpetrators and women are the victims overlooking the growing literature on male victimization forms of sexual assault. Since its highly likely that victims will be raped by men according to their theory, males have the tendency to associate rape with masculinity and power [15]. Most males who have been sexually assaulted will not admit to sexual assault as they deem victimization as a form of weakness. Women and girls on the other hand to some extent will report sexual assault but only in situations where they are absolutely sure they will get help [16].

1.2 Statement of the problem

Preventing sexual assault is necessary to support human rights and global development goals, including the Sustainable Development Goals (SDGs). Sexual assault is a significant problem in many parts of the world, including Kenya. Understanding the dynamics around rape perpetration and victimization is key to addressing this public health problem. By being able to predict what will happen on rape victimization if factors are kept constant, it is possible to design both effective and cost effective interventions to mitigate this problem. Nairobi's 4 largest informal settlements are no exception to this vice. With every 1 girl in 4 being sexually assaulted annually, there is need for the understanding of the perpetration dynamics within these slums to help mitigate this problem that affects adolescent girls. Even though there are several interventions in these areas to mitigate rape, there is little information that can be scientifically used to inform the design of even better interventions by defined by being robust enough to classify or determine which individuals are at the highest risk of being sexually assaulted. A lot of intervention currently being rolled are aftercare based, basically reaching out to the girls once they have been sexually assaulted. This is a clear indication that there is limited understanding of rape perpetration and victimization patterns to inform the design of effective preventative interventions. There is need to be able to predict the vulnerability levels of girls living in these environment.

1.3 Study Objectives

To determine victimization dynamics in the 4 slums of Nairobi using existing rape perpetration data from the slums.

1.3.1 Specific objectives

- To come up with a classification model for rape with Age, Slums and Grade as predictors
- Determine which of these 3 variables contribute most to the likelihood of getting sexually assaulted.
- To determine who are the perpetrators and who are their targets by age, grade and slum.

1.4 Research Questions

1. Can existing data on rape perpetration in 4 slums of Nairobi be used to determine rape victimization dynamics in these slums amongst high school girls?
2. Do Age, Grade and slum determine the likelihood of a girl being sexually assaulted?
3. Who are the perpetrators and who are their targets by age, grade and slum?

1.5 Justification

Sexual assault is a huge public health problem in Nairobi's 4 largest slums. Studies show that 1 in every 4 girls in these slums has been sexually assaulted. There is therefore need to understand the dynamics of rape perpetration and victimization in these demography to help inform mitigation and policy. This study will be using data collected over the last 3 years from quasi experimental studies centered on a rape prevention program in these slums to help predict/ model rape perpetration dynamics on adolescents. Even though there are a lot of research activities in slums around sexual assault, most are centered on primary data collection which are often very expensive and time consuming. With existing data from these slums, it is possible to statistically predict victimization dynamics within the slums that would then inform the design of effective interventions against this social problem. Data collected from these slums over the last three years by a nonprofit called Ujamaa Africa can go a long way in ensuring that this is achievable. With a clear model of assessing individuals risk factors

CHAPTER 2

INTRODUCTION

2.1 Introduction

Understanding the dynamics around rape perpetration and victimization is key to addressing this public health problem. By being able to predict what will happen if factors are kept constant, it is possible to design very effective interventions to mitigate this problem. There are a number of studies that have looked into perpetration dynamics. A lot of these studies have focused on global perspective around rape risk factors as opposed to the African context which is relatively unique and is driven by a host of different factors from those in the West or Asia.

2.2 Literature Review

Studies have identified rape risk factors that cuts across different countries and different continents. Despite the fact that many countries experiencing elevated levels of rape are geographically distinct, sexual assault risk factors are more or less the same across these groups [2, 3, 5]. Studies with strong statistical approaches have over the years investigated just how significant or not some of these factors are across different contexts. Employing logistics regression approaches, studies have demonstrated that age and sex are some of the most important risk factors of rape [7]. In a study conducted in England, multivariate logistic regression identified sex and age significant demographic risk factors for rape [14] and that most of the perpetration is carried out by men. In South Africa for instance, a study showed that 8.8% of men had at one point in their lives participated sexual assault [15].The

same study also demonstrated that the risk factors for single perpetrator rape do not differ with those of multiple perpetrator rape. The study and many others demonstrate that the risk of sexual assault is significantly higher on girls than for boys and are driven by specific factors that could vary from one context to another.

2.2.1 Rape risk factors in Africa

One of the key questions that researchers have been trying over the years to answer is that of specific risk factors for adolescent girls in different context. Even though these risk factors are relatively the same across different countries there are however very specific factors that drive perpetration and victimization in each context [7, 8, 15]. In most African countries affected by high incidences of rape there are generally similar factors that drive sexual assault directed towards adolescents. These factors can be generalized to high social status or family background (well educated parents), low self-esteem, engagement in anti-social behavior such as bullying at school and substance abuse [17]. Also, engagement in activities that show an exaggerated and dominant sexuality (Negative masculinity), having more than one partner and engagement in regular transactional sex [18]. These are factors that drive perpetration of rape. Rape supportive notions over the years have formed the basis for victimization.

2.2.2 Rape victimization scales

Several scales that are used to measure masculinity ideologies beliefs have been postulated and hypothesized to support sexual assault and other forms of male dominance [15]. These theories have been generated via feministic socio-cultural models of sexual assault and other forms of male domination, their early successes in accounting for rape victimization

and perpetration, their direct correlation to the many interventions and their ease of administering to students [3,16]. This beliefs build from those that are associated with adversarial nature in which men relate to women, stereotypical ideologies about sexuality and sexual roles, negative pervasive notions on femininity and masculinity and sexual assault myths that encourage sexual aggressiveness.

Murnen et al. (2002) carried out a meta-analysis looking at 11 different assessors of masculinity ideologies in 39 study sites. In their study, they found out even though all except one measure of masculinity ideology was significantly associated with rape, the strongest contributors came from hostile negative hyper-masculinity and masculinity, both of these measure hostile beliefs around girls and women, the intense desire and need to be in control, and the general acceptance of domination and violence against women [16, 27]. Assessors of simplistic gender adherence which excluded hostile behavior and the general acceptance of aggressive aspects were not very strong predictors of sexual aggression [5].

2.2.3 Non-partner rape risk

In a recent 38-site cross-cultural study, in 2007, Hines and his group looked at the role of aggressive sexual beliefs, women's status in the society, and also prior rape victimization as risk factors for sexual coerciveness against both men and women. In her study just like the South African study, she found that both aggressive attitudes around relationships and also prior sexual victimization related to sexual aggressiveness, across cultures, in both genders across samples non-criminal and criminal in nature [17, 18]. Studies that have recently used Multidimensional Inventory of Development, Sex, and Aggression (MIDSA)

found out that negative gender masculinity and hostile behavior towards women effectively differentiate sexual offenders in criminal samples as opposed to non-criminal ones (MIDSA, 2008).

Studies have demonstrated age of an individual as risk factor for sexual victimization. Perpetrators especially those close to their victims usually target individuals who would give the least resistance [17]. In a study conducted in the DRC, (Amber et al, 2011), school going or girls of school going age were targeted the most. Using a house hold survey that involved 3436 women and girls, they were able to demonstrate age as a desirability factor in sexual assault using multivariate logistics regression. The study suggested that girls between the ages of 15-25 years old were most likely to be sexually assaulted. The results from this study is consistent with those from many other parts of Africa and the world [23]. In low income countries especially around demographics living in informal settlements, girls of school going age are the highest targets of sexual assault [3, 37].

2.2.4 Adolescent girls victimization risk

In a study carried out in Nairobi's Slum in 2011, (Sinclair et al, 2011), 1 in every 4 girls was sexually assaulted in the year prior to the survey. These were girls who attended secondary schools within the slums. Majority of these rapes took place along the way to school with a significant number of the rapes occurring in school or other learning environments [3]. Even though both young and older girls use these paths, there are fewer reported cases of these other age groups being sexually assaulted. The quasi experimental study also demonstrated that women and girls were also highly likely to be sexually assaulted by individuals close to them, usually relatives and close guardians. Data from around the world suggests that young girls and women are usually at more heightened risks

of sexual assault than older women [19] According to information gathered by rape crisis centers and the justice systems in Malaysia, Chile, Mexico, Papua New Guinea, the United States and Peru, between a third and two thirds of all survivors of sexual assault were aged 15 years or less [13, 23, 33]. Certain forms of sexual assault, for example, are usually closely associated with the young, in particular violence taking place in learning institutions and colleges.

As girls transition into puberty and become more sexually active, they become more sexually desirable to men [16]. School activities and learning environment take these newly sexually active individuals from their protective home environments into the hunting arenas of potential perpetrators. As they get more educated on sexuality other related things, women tend to be able to make better decisions around their sexuality and judgment around these issues [18].

Studies have demonstrated that women and girls are at a heightened risk of sexual assault, like they are of physical assault by men or partners, when they become more educated and thus more empowered. Girls and women with little education have been shown through a South African national survey to be much less likely to be sexually assaulted compared to those with more education [16, 18]. In a study carried out in Zimbabwe, working class women were more likely to report sexual assault meted on them by a stranger or even a spouse [19]. The most likely explanation to this phenomenon is that with empowerment, women are more likely to resist and question patriarchal societies that many of them live

in. The level of education with respect to grade reached in school is an important factor that determines risk of sexual assault.

2.2.5 Socio-economic status and sexual assault

The economic status of an individual determines their likelihood of getting sexually assaulted, this affects both the victim and the offender [17]. Several studies have demonstrated that the link between perpetration of sexual violence and poverty is mediated through various forms of masculinity identity. In a study conducted in 2001, it was argued that extended periods of idleness, unemployment or lack of jobs, economic difference between the rich and the poor, insufficient institutional support from the police and the judicial system, and the general acceptance and tolerance of rape and other forms of sexual assault contributes most to the perpetration of rape by men [16].

Rape and other forms of sexual assaults are usually associated with external environments away from the protective home environment, contrary to this, a lot of sexual assault happens within the confines of these protective environments. According to Marrel and his group in a study conducted in 2001, the social environment that drives rape is more important than the physical environment. When people are socialized to trust those close to them to protect them, they are exposed to the risk of being assaulted by these very people. The male entitlement phenomena that comes with most patriarchal societies allow for this negative environments to thrive [16, 17].

At the national and the international levels, existing policies contribute to the rape. This is not limited to the not so conducive economic environment that the national governments create. Economic melt-down and global financial crisis continues to feed the existence of these high risk areas such as slums [19].

The key question given all this risk factors would be to find a combination of the best factors that would separate best girls into two distinct groups. Girls likely to be sexually assaulted and those that are unlikely to suffer the same. Various mathematical models have been used to achieve this.

2.2.6 Mathematical models for risk assessment

Various statistical procedures have been used to analyze causal aspects of situations in behavior science in various studies. The main mathematical components used in these kinds of analysis include Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA) and Independent Component Discriminant (ICD) analysis approaches are usually used or applied in situations where we are dealing with multivariate situations where normally data collection is done at a single data point in time [20]. DA can also be used in discussion around studies that focus on repeated measures approaches, in these cases study participants or respondents provide responses at two or more predetermined data points. Repeated measures approaches usually come up in many research areas not limited to social and behavioral science disciplines.

Linear discriminant analysis has been used successfully to carry predictive analysis on variables [20]. LDA simply looks for linear combination of variable that best separates two groups. Using continuous predictors and categorical dependent variable, DA can has been used in both social and behavior research to better understand causal pathways of various phenomena under investigation. Allocatory DA techniques are extremely powerful tools for designing statistical classification models and algorithms [24]. These techniques are frequently used in pattern recognition and data analysis. In the case of two population and

assuming that all parameters are known, the best statistical classification law that minimizes the probability of misclassification (PMC) is the Bayes rule [20]. In designing sample based discriminant functions for statistical analysis, there are two main approaches that are parametric in nature. The first one is the plug-in approach that assumes the class densities are known while the second one is the Bayesian approach in where prior density on the parameter vector is assumed to be known [36]. For this paper will use the Bayes approach of working with prior density to come up with a classification criteria for rape victimization risk for a select group of girls residing in the slums. For this study we will be employing the principles of Discriminant analysis and Logistics regression to analyze predictor of sexual assault on a population of girls residing in Nairobi's Informal settlements. These two approaches fit our data best for predictive analysis.

2.3 Research Gaps

There is however very little data or research on what risk factors drive victimization. Most studies will look into gender norms, men's perspective of women and the role of substance abuse on the perpetration of rape. This gap in research is more profound in populations living in informal urban settlements in major towns and cities in Africa. The few studies that have looked into the risk factors have come up with different results. One study conducted in South Africa found that their logistics regression model showed that the risk factors for sexual assault were, childhood experience of violence, alcohol use, corporal punishment at home and being negative about school. Very few studies have looked at the combined effects of the various postulated factors. Predictors of rape victimization has always been looked at individually without using statistical approaches that bring all the predictors together to look at the effect they have on victimization, especially in African rural and urban setups. With the need to design more targeted and age appropriate interventions, there is need to understand the causal pathways and risk factors of sexual assault in informal settlements. There is lack of robust means of classifying adolescent based on risk factors that might help in more targeted intervention design.

With many of studies investigating the drivers for perpetration there is limited research on the factors or drivers of victimization. What appeals most to the rapist?

CHAPTER 3

Methodology

3.1 Research Design and Methodology

Participants were a prospective cohort of 2383 high school girls 14-20 years of age, who attended one of 35 high schools selected through school convenience sampling from 4 informal settlements of Nairobi that is Huruma/Mathare, Dandora, Kibera, and Mukuru. These areas were selected because of their high rates of crime and having not participated in the previous studies conducted by Ujamaa Africa.

3.2 Population and Sampling

Neighborhoods were purposively assigned to the intervention group or Standard of Care SOC group randomizing participants individually could not be possible given that the participants were often neighbors and friends. Participation was on voluntary basis. The participants were given an opportunity to review and sign designed informed consent forms. No students declined to participate. All the participants came from extremely poor homes, usually from houses and homes constructed from found material. Most of their homes either lacked a sewer or fresh water. Participants from these schools came from different tribes. The most predominant ethnic group were Luo, Luhya and Kikuyu, Table 1 below shows the demographic characteristics of these slums.

For this paper the analysis will focus on the baseline data collected before the commencement of a self-defense and empowerment training for the participants by Ujamaa Africa, an NGO that works to prevent violence against women and girls. All the girls who participated in the baseline surveyed were included in the study. Students enrolled in each

school ranged from 10-204 students, (90.3%) of the schools, that is 28 schools an

enrollment rate of 100 students and less. This study underwent review and approval by The

National Commission for Science, Technology and Innovation (NACOSTI)

Table 1 **Characteristics of study slums and schools (Similarities across slums)**

	Huruma	*Kibera*	*Dandora*	*Mukuru Kwa Reuben*
Pop (2009 Census)	106,913	170,070	135,000	100,000
Size (Square km)	1.4	2.4	1.5	1.5
Pop. Density (people/km^2)	76,366	70, 862	90,000	66,666
Monthly Income	Ksh 7,000 (USD 85)	Ksh 5,000 (USD 60)	Ksh 6,500 (USD 80)	Ksh 4,400 (USD 50)
Single room monthly rent	Ksh 2000 (USD 24)	KSh 800 (USD 9)	Ksh 2000 (USD 24)	Ksh 1,400 (USD17)
Water Price (20 litres)	KSh 10	KSh10	KSh 8	KSh 2
School Building Material	Iron sheet/ Stone	Iron sheet	Iron sheet	Stone/ Iron sheet
School Floor	Cement	Cement/ Earth	Cement	Cement/ Earth
Trained: Untrained teacher ratio	1:2	7:5	3:4	1:5
Teacher: Student Ratio	1:25	1:20	1:16	1:15

National exam mean score 2011*	C+	D+	D	C-

.

Figure 1. Distribution of slums by students sampled
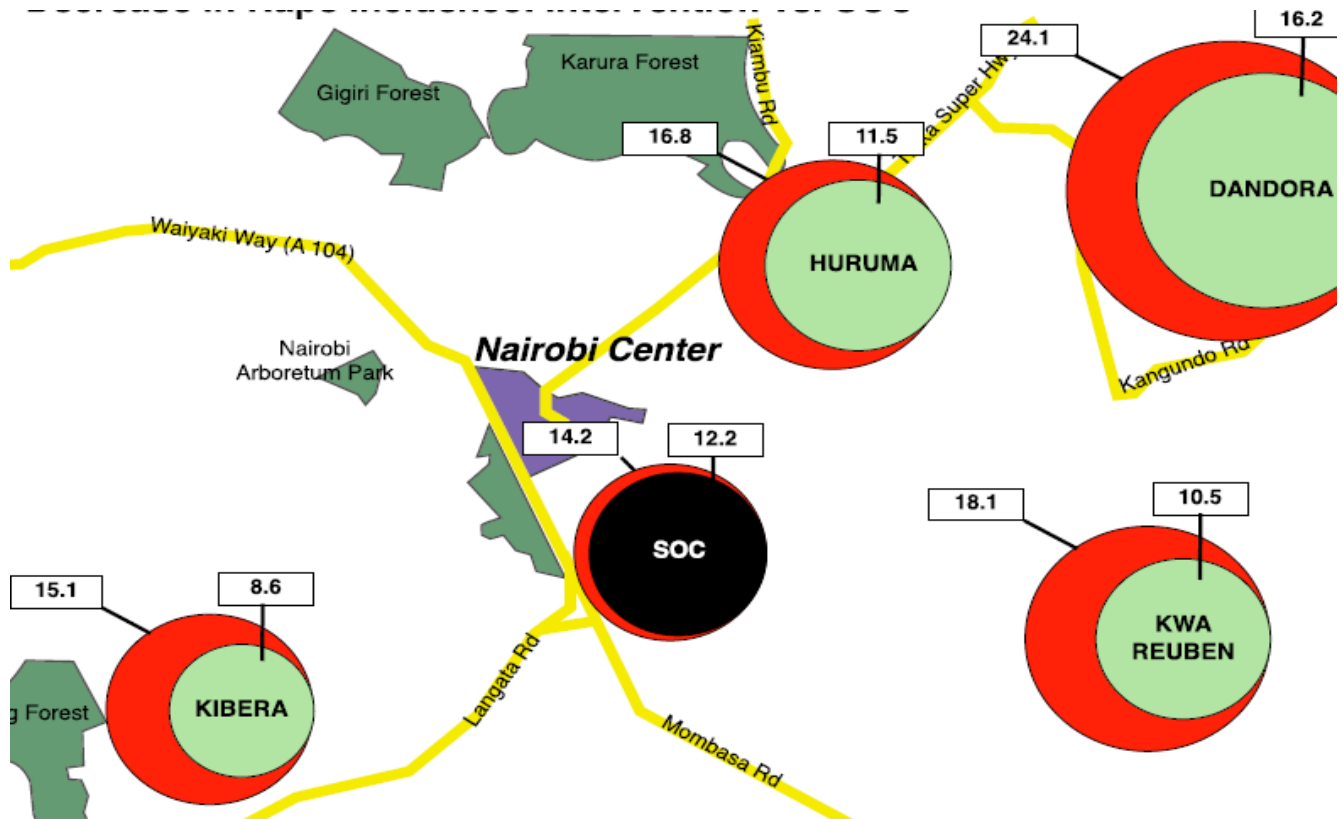
3.3 Data Collection

The main outcome measures for the overall study were a drop in the annual incidence of rape among the girls enrolled in the program. In this survey, 'rape' or 'sexual assault' was defined as "forced penetration of the Vagina, anus, or mouth by penis, an object or any other body part".

Sexual harassment on the other hand was defined as "unwanted gestures, comments or whistles that came with sexual intentions", Data on sexual assault, demographics, and risk factors were collected via surveys, which were Structured and close-ended in English . The data was anonymously collected through the Ballot Box Method (BBM) [19] by trained data collectors. Surveys were administered before the first class (baseline).

Data was collected in schools premises after school sessions ended (3-6pm) by qualified data collectors under the supervision of Ujamaa Africa's Research department and the schools' guidance and counselling teachers. Participation was voluntary and students and students who did not feel like participating in the survey were free to leave the survey rooms. The survey were self-administered with a data collector reading the questions in English and Swahili to all the students who in turn filled out the questionnaires. The respondents were divided into different classes and spaced 1 meter apart from each other as they filled in the survey.

For this study however the following variables collected from the study formed the basis for the analysis, that is data on, *'last one year incidences of rape'*, '*Who raped you?'*, *'Age of respondent'*, *'Slum from where the girls come from'* and *'girl's grade in school'*. The survey used to collect this data is attached as appendix 1

3.4 Data Entry

The anonymously collected data was brought to a central office at the headquarters of Ujamaa Africa, serialed collated and processed for data entry, obvious inconsistencies were resolved by the data team.

Data was double entered into SPSS; inconsistencies were resolved by a third party and entered.

Statistical analysis for rape was made using incidence rate ratios 95% Poisson Confidence Intervals (CIs). Statistical analysis for rape and harassment disclosure, annual incidence of rape lifetime incidence of rape, and age were done using Fisher's exact tests and chi square tests.

Level of significance was set at 0.05. Data analysis was done with the aid of SPSS 20.0 (IBM) and R statistical software.

3.5 Data summary

3.5.1 Sample size

Out of the 2383 girl's sample, 672 were sexually assaulted in the last one year. Fig 1 - 6 shows the distribution of these incidences by slum, grade, age and perpetrators respectively. Each participant provided data on the above predictors

Figure 2: Students sampled by Slum

	Dandora	Huruma	Kibera	Mukuru
Series1	639	958	630	156
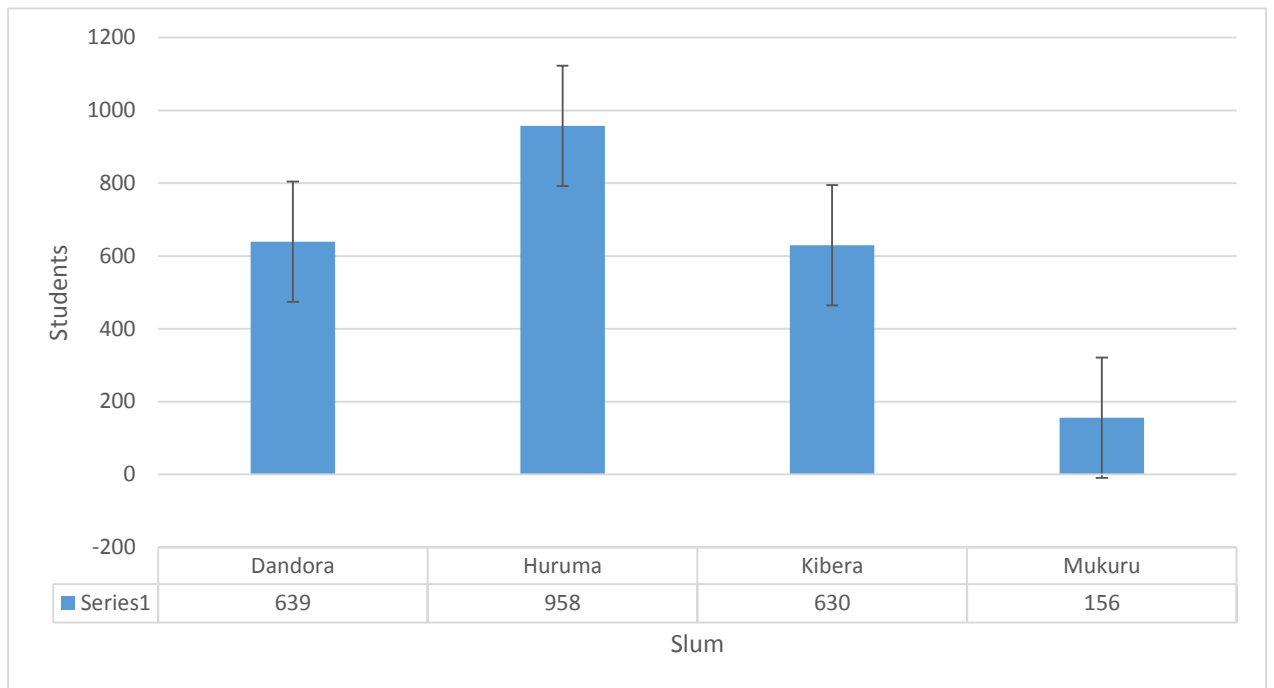
Slum

A total of 2383 girls participated in the study. Huruma posted the highest number of student with a total of 958 girls filling out the questionnaires. Mukuru had the least number of participants with 156 surveys filled.

3.5.2 Rape by slum

Figure 3. Rape incidences by slum

	Dandora	Huruma	Kibera	Mukuru
■ Series1	259	285	100	28
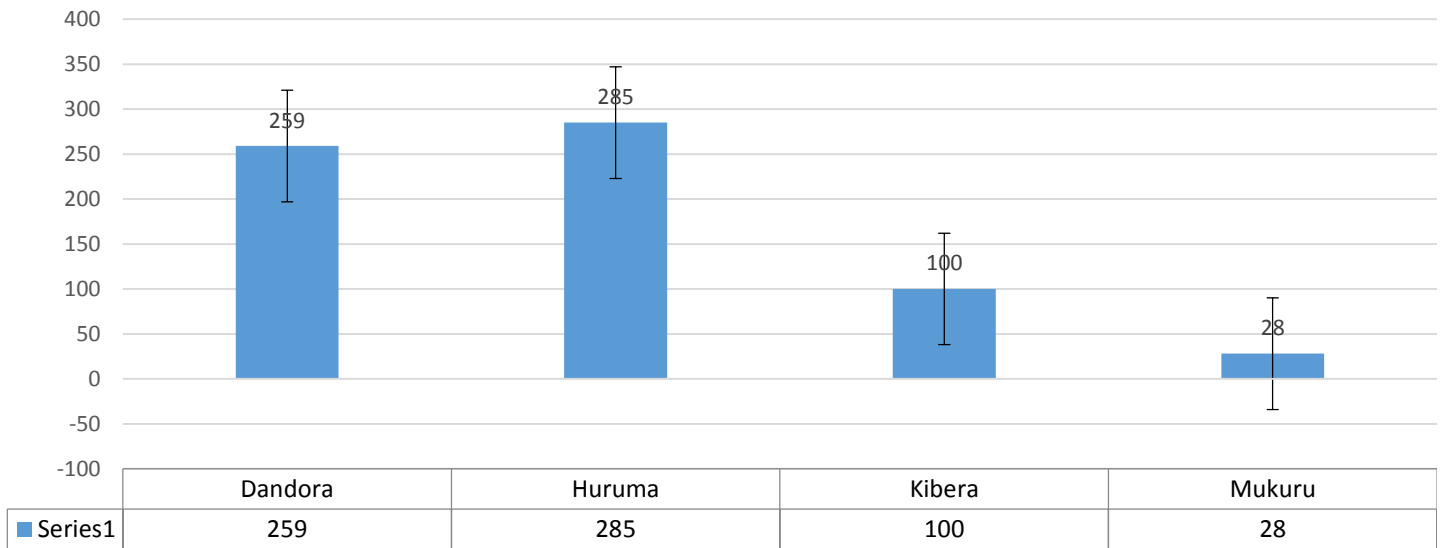
Dandora and Huruma slum bore the brunt of the sexual assaults. Despite Kibera and Mukuru being the largest slums, they recorded the lowest numbers of rapes. Huruma slum had the highest incidence with 285 cases. Mukuru slum had the lowest incidence with only 28 rape cases

3.5.3 Rape by grade

Figure 4. Rape by Grade

	Form1	Form2	Form3	Form4
■ Series1	55	333	144	140
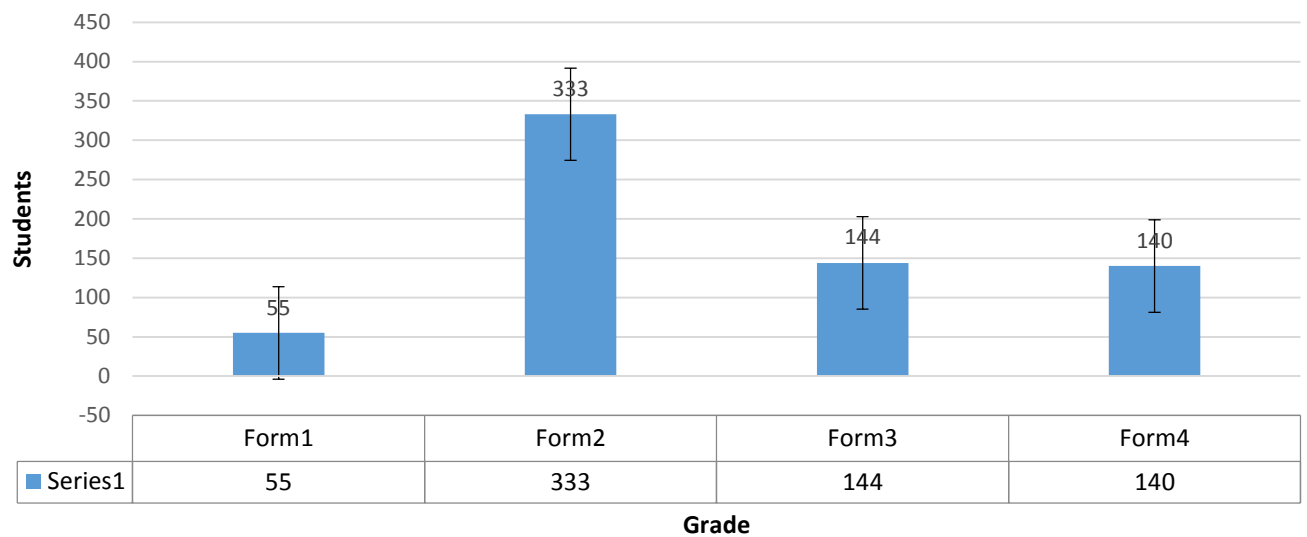
Grade

Students in Form 2 appear to be the most sexually assaulted of all girls. Form 2 as a grade had 333 rapes, Form 1 had the least rapes with 55 reported cases. Form 3 and 4 had relatively the same number of rapes 144 and 140 respectively.

3.5.4 Rape by Age

Figure 5. Sexually assaulted girls by Age

Series1	14 year olds	15 year olds	16 year olds	17 year olds	18 year olds	19 year olds	20 year olds
	10	22	76	235	181	143	5

Age
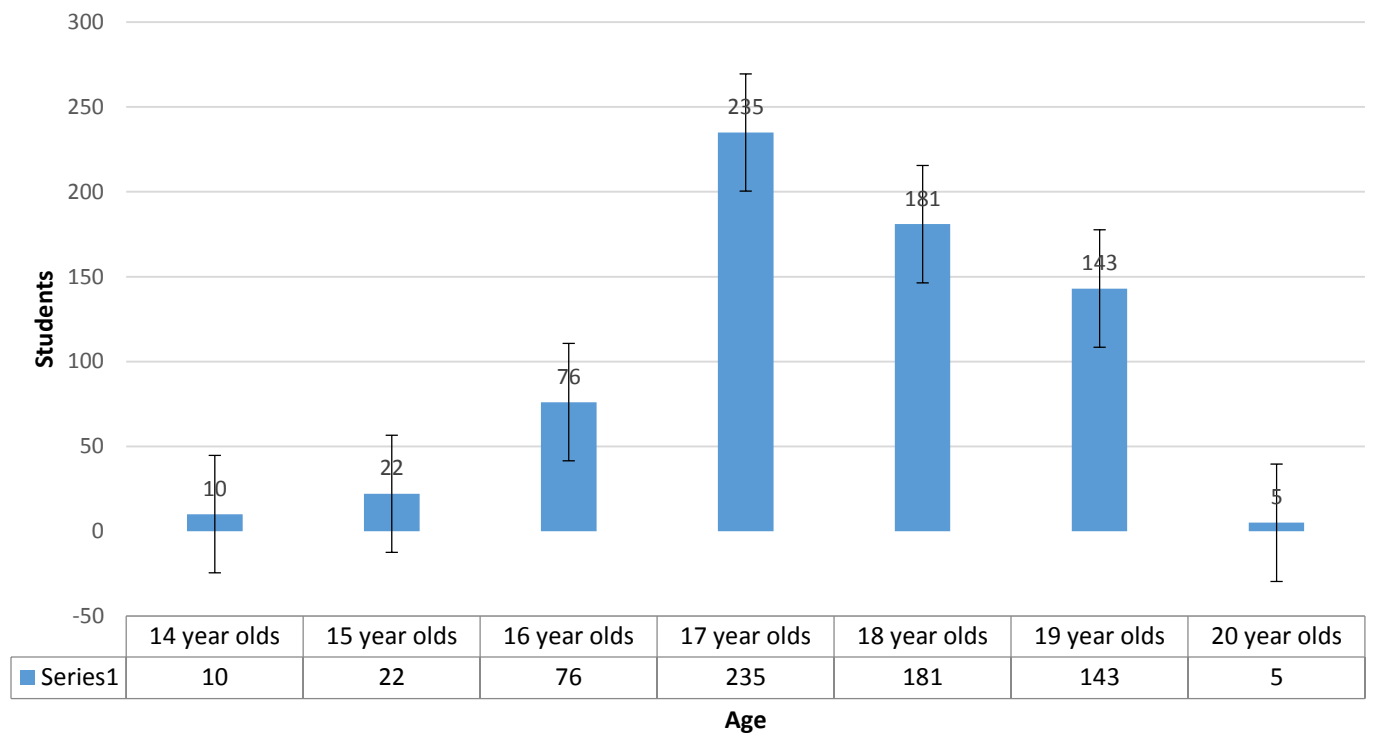
With an age distribution of between 14 -20 years old, 17 year old recorded the highest number of rapes, 235. 20 year olds recorded the list number of rapes at 5 incidences.

3.5.5 Rape by Perpetrators

Figure 6. Rape by perpetrators (Who raped)

Perpetrator	Boyfriend	Doctor	Friend	Gangster/stsranger	Neighbor	Pastor	Police	Relative	Teacher
■ Series1	121	3	56	67	55	14	5	32	7
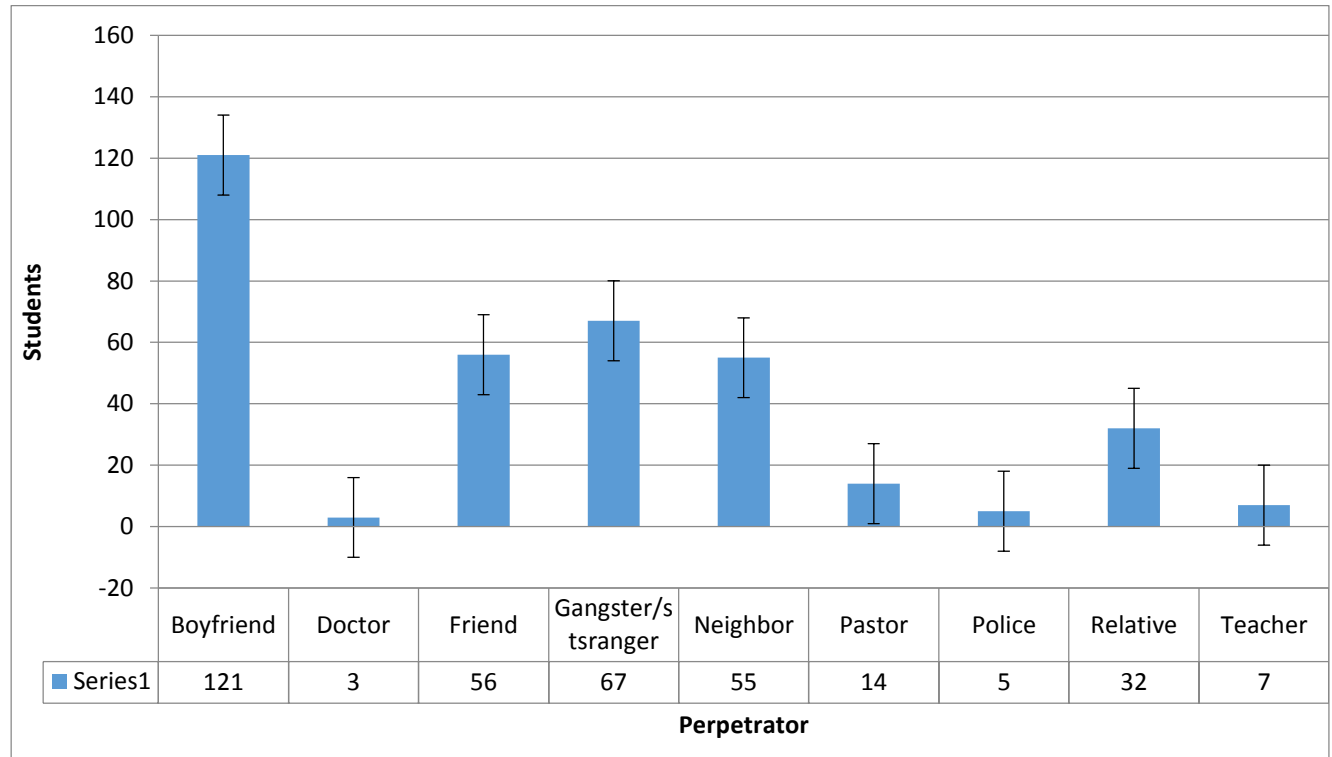
Majority of the rapes were perpetrated by intimate partners (Boyfriends), this accounted for 121 cases. Majority of the girls were sexually assaulted by individuals close to them. Friends 56, Neighbor, 55 and relative 32. The police and doctors recorded the lowest incidences with 5 and 3 incidences respectively

CHAPTER 4

4.0 Statistical Analysis and Results

For categorical data, data was analyzed using chi square and fishers exact test. For continuous variables independent t-tests were used.

General linear model to determine variables contributing most to rape

We went further did regression analysis using the same predictors, that is *Age of respondent (Age), 'Slum from where the girl come from' (Slum) and 'girl's grade in school' (Grade)* to determine which of these predictors contributes most to either being raped or not.

Table 2. Predictors and levels

Predictor	Levels			
Grade	Form 1	Form 2	Form 3	Form 4
Slum	Dandora	Mukuru	Kibera	Huruma
Age	Continuous 14 -20 years			

Rape Versus slum regression (Contribution of individual slums)

With Dandora slum as the reference slum

Call: Glm (formula = Raped ~ slum, family = "binomial", data = dataset)

Deviance Residuals:
```
  Min        1Q      Median       3Q       Max
-1.0195    -0.8404   -0.5879    1.3439    1.9186
```

Coefficients				
	Estimate	**Std. Error**	**z value**	**Pr (>|z|)**
(Intercept)	-0.38334	0.08058	-4.758	1.96e-06 ***
SlumHuruma	-0.47591	0.10718	-4.440	8.98e-06 ***
SlumKibera	-1.28436	0.13557	-9.474	< 2e-16 ***
SlumMukuru	1.13648	0.22365	-5.082	3.74e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null deviance: 2835.0 on 2382 degrees of freedom
Residual deviance: 2727.3 on 2379 degrees of freedom
AIC: 2735.3

Number of Fisher Scoring iterations: 4

Rape Versus Grade regression (Contribution of Grades)

With Form One as the reference Grade

Call:

Glm (formula = Raped ~ Grade, family = "binomial", data = dataset)

Deviance Residuals:

   Min   1Q Median   3Q   Max

-0.9897 -0.8125 -0.7115   1.3775   1.9675

Coefficients				
	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	-1.7796	0.1458	-12.208	< 2e-16 ***
GradeForm 2	1.3205	0.1617	8.166	3.19e-16 ***
GradeForm 3	0.5348	0.1738	3.078	0.00209 **
GradeForm 4	0.8407	0.1766	4.760	1.93e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 2835.0 on 2382 degrees of freedom

Residual deviance: 2738.7 on 2379 degrees of freedom

AIC: 2746.7

Number of Fisher Scoring iterations: 4

Rape Versus Age regression (Contribution of Age)

Call:

Glm (formula = Raped ~ Age, family = "binomial", data = dataset)

Deviance Residuals:

  Min   1Q Median   3Q   Max

-1.0332 -0.8658 -0.7889 1.4265 1.9180

Coefficients				
	Estimate	**Std. Error**	**z value**	**Pr (>|z|)**
(Intercept)	-4.73933	0.66677	-7.108	1.18e-12 ***
Age	0.21950	0.03821	5.744	9.24e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null deviance: 2835.0 on 2382 degrees of freedom

Residual deviance: 2800.8 on 2381 degrees of freedom

AIC: 2804.8

Number of Fisher Scoring iterations: 4

Combined GLM analysis using R.

Rape Versus Age, slum, and grade regression

Call:

(Raped ~ Age + slum + Grade, family = "binomial")

Coefficients				
	Estimate	**Std. Error**	**z value**	**Pr (>|z|)**
(Intercept)	-5.44623	0.81851	6.654	2.86e-11 ***
Age	0.26587	0.04931	5.392	6.97e-08 ***
SlumHuruma	-0.41122	0.11060	-3.718	0.000201 ***
SlumKibera	-1.09916	0.13945	7.882	3.22e-15 ***
SlumMukuru	-0.95410	0.22714	-4.200	2.66e-05 ***
GradeForm 2	0.86483	0.17133	5.048	4.47e-07 ***
GradeForm 3	0.05125	0.18961	0.270	0.786936
GradeForm 4	0.06444	0.21111	0.305	0.760195

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 2835 on 2382 degrees of freedom

Residual deviance: 2632 on 2375 degrees of freedom

AIC: 2648

Number of Fisher Scoring iterations: 4

Two of these predictors contributed most to either being raped or not. The slum from where the girls comes from and the girl's age contributes most to the likelihood of getting raped.

Linear Discriminant Analysis (LDA) for prediction of raped or not raped

Using the dataset, LDA was performed to determine the accuracy of predicting whether a girl is raped or not. With "RAPE" (Yes/No) as the dependent variable *Age of respondent (Age), 'Slum from where the girl come from' (Slum) and 'girl's grade in school' (Grade)* were used as predictors for the LDA analysis. Table 2 shows the predictors and their respective levels.

Table 3. Predictors and levels

Predictor	Levels			
Grade	Form 1	Form 2	Form 3	Form 4
Slum	Dandora	Mukuru	Kibera	Huruma
Age	Continuous 14 -20 years			

The Coefficients of LDA, that is the Linear Discriminant Function table usually interprets the Fisher's theory, therefore it is only useful when a linear model is chosen for the Discriminant Function [20]. Discriminant analysis is based upon the science of coming up with a linear combination of predictors which separates best two classes (targets) Fisher defined the following score function

$$Z = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_d x_d$$

$$S(\beta) = \frac{\beta^T \mu_1 - \beta^T \mu_2}{\beta^T C \beta} \qquad \textit{Score function}$$

$$\downarrow$$

$$S(\beta) = \frac{\overline{Z}_1 - \overline{Z}_2}{\text{Variance of } Z \text{ within groups}}$$

Estimating of the linear coefficients that maximizes the score takes the following form.

$$\beta = C^{-1}(\mu_1 - \mu_2) \qquad \textit{Model coefficients}$$

$$C = \frac{1}{n_1 + n_2}(n_1 C_1 + n_2 C_2) \qquad \textit{Pooled covariance matrix}$$

Where:

β : *Linear model coefficients*

C_1, C_2 : *Covariance matrices*

μ_1, μ_2 : *Mean vectors*

The linear discriminant functions, also called "classification functions", for each observation, take the following form

$$C_k = C_{k0} + C_{k1} X_1 + C_{k2} X_2 + \ldots + C_{km} X_m$$

where

- C_k is the classification score for group k

- $C's$ are the coefficients in table

For each individual observation, the method compute the score for each distinct using the coefficients following equation (1). The entry is then assigned to the side with highest score. Besides this, the calculated coefficients help in deciding which of the many predictors has more effect in classification.

If we let *yij* be a *p × 1* vector of observed measurements on p variables in a given dataset, where membership of group is known for the ith project participant *(i = 1, ..., nj)* in the jth group (j = 1, 2). This works for two group designsThe assumption here us *yij* ~ *Nq (μj, Σj)*, where *μj* population mean vector, *Σj* are the covariance matrix for the *jth* group and that are to be estimated respectively by $\mu\hat{j}$ and $\Sigma\hat{j}$,

The Discriminant Analysis rule takes the form below. The *ith* participant is assigned to group 1 using the equation below.

$$\lambda\left(\mathbf{y}_{ij}\right) = \left[\mathbf{y}_{ij} - \frac{1}{2}(\widehat{\mu}_1 + \widehat{\mu}_2)\right]^{\mathbf{T}} \widehat{\mathbf{a}} > \ln\left(\frac{\widehat{\pi}_2}{\widehat{\pi}_1}\right)$$

The linear DA classification rule is: Assign the *ijth* study participant to group 1 if. The else statement assigns the participant to group 2. In the above Equation T refers to the Transpose

In the above equation, the estimater for the linear discriminat function, **a**

$$\widehat{\mathbf{a}} = \widehat{\Sigma}^{-1}\left(\widehat{\mu}_1 - \widehat{\mu}_2\right)$$

P a g e | 36

Where

$$\widehat{\sum} = \frac{(n_1 - 1)\hat{\sum}_1 + (n_2 - 1)\hat{\sum}_2}{n_1 + n_2 - 2}$$

π_1 and π_2 in the equation are priori probabilities that the participant belong to group 1 and 2, respectively. This is usually estimated by.

$$\widehat{\pi}_j = \frac{n_j}{N}$$

Where $N = n_1 + n_2$

The relative importance of predictors for a discriminant analysis to separate between two groups and can measured by standardized coefficients by using the magnitude of the absolute value.

The ultimate goal of LDA is to separate two groups and determine the accuracy of the separation. DA does this by computing the Misclassification Error usually abbreviated as (MER)

This is basically the probability that an individual or rather a participants is incorrectly classified into either groups. The Apparent error rate (APER) is calculated as below

$$APER = \frac{N - n_{11} - n_{22}}{N}$$

n_{11} and n_{22} are absolute numbers of study respondents or participants who have been correctly assigned to either group 1 or 2 respectively.

In LDA, the discriminant model produced assumes a Multivariate Normality, Equality of variance-covariance within group and Low multi collinearity of the variables. When high multicollinearity among two or more variables is present, the discriminant function coefficients will not reliably predict group membership

Figure 7 below shows the results from the LDA Model using R statistical software.

Call:

Ldaout <- lda (Raped~Age+Huruma+Mukuru+Kibera+Form2+Form3+Form4, dataset)

π_1- *Prior Probability of group* ***Yes*** *(Raped)*   *= 0.2819975*

π_2- *Prior Probability of group* ***No*** *(Not Raped) = 0.7180025*

Figure 7: Coefficients of linear discriminants

Coefficients of linear discriminants	
	LD1
Age	0.4003174
Huruma	-0.5987363
Mukuru	-1.014703
Kibera	-1.1680299
Form2	1.210587
Form3	-0.4147623
Form4	-0.7947091

Figure 8: LDA Classification plots
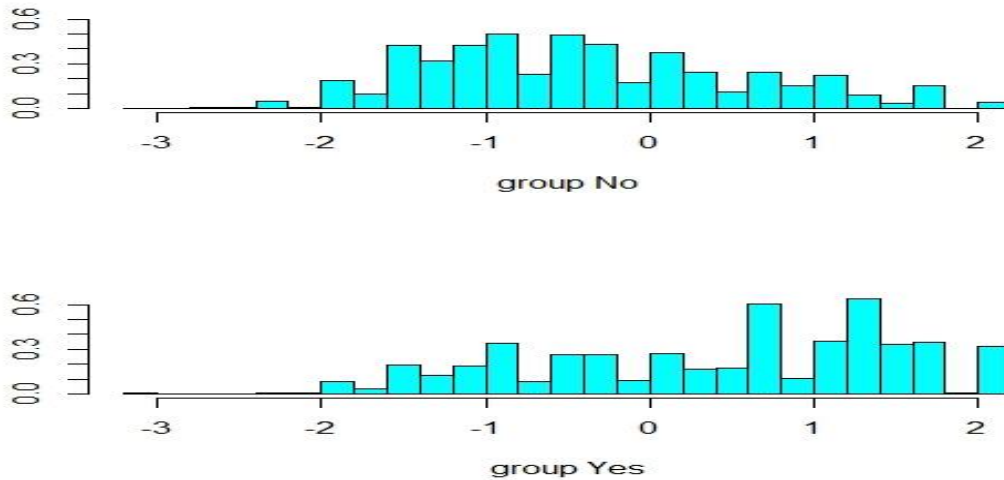
Table 4: Group means

Group Means							
	Age	Huruma	Mukuru	kibera	Form 2	Form 3	Form 4
No	17.16657	0.3933372	0.07481005	0.3097604	0.3290473	0.2822911	0.198130
Yes	17.49405	0.4241071	0.04166667	0.1488095	0.6398810	0.1636905	0.1145833

Prediction using the LDA model's Coefficients

Using these coefficients, the model was tested on the same dataset used to generate the model to see how it performs in classifying individuals into either Raped or Not raped categories, the figure below summarizes the results of the prediction, comparing the actual dataset to the predicted R output. Using these predictors the LDA model created was able to categorize respondents into raped or not raped categories with 76.62% accuracy

Actual data vs Predicted

Raped from actual data	
No	**Yes**
1711	672
Actual data	

	No	Yes
No	**1604**	450
Yes	107	**222**
Prediction using LDA		

From the tables above the LDA model was able to correctly classify 1604 out of the 1711 Nos. 222 out of the 672 Yeses were also correctly classified. This gives a combined 76.62% accuracy in classification.

$$APER = \frac{N - n_{11} - n_{22}}{N}$$

APER = 23.38% (Miss-classification rate)

The model will misclassify rape 23.38% of the times.

CHAPTER 5

Discussion and Conclusion

Incidence of sexual assault

In this group of impoverished Kenyan adolescent girls, the annual incidence of sexual assaults is 28.19%. This rate is significantly higher than those reported for the Nairobi region by the Kenya police, where the annual incidence of 5.5% in 2011, or the African Population and Health Research Center, at a lifetime incidence of 12.1%, also the Kenya Demographic and Health Survey, showed a 14.5% lifetime risk [21] The significantly higher rates may result from the demographics studied, young residents of high-risk areas, or from the confidential and anonymous method used to collect data. Nonetheless, these results are consistent with other reports from other parts of Africa, where 23% of school going girls between the ages 13 and 15 years experienced sexual violence [21]. Generally a girl in a typical Nairobi slum stands a 28% chance of getting sexually assaulted.

Predicting of rape or no rape

Consistent with the findings of the Linear Discriminant Analysis in this study, Age, grade and slum can be used to determine the risks or predict the chances of getting sexually assaulted. From the dataset of 2383 students surveyed at baseline it was possible to correctly place individuals in their respective category with respect to either being raped or not with a 76.62% accuracy. Studies from different settings have shown that these individual predictors are a very big component of characterizing sexual assault risk. In LDA, Predictors are given higher priority in a hierarchical analysis giving essentially what

would be a discriminate function analysis with covariates. It does this by creating discriminate functions and each is assessed for significance.

Using the three predictors of rape (Age, Slum and Grade) in this manner the LDA creates a model that assigns an individual to either being raped or not. LDA works best with continuous predictors and categorical Y variables, the resulting model from this analysis can gives 76.62% accuracy in prediction. This can be explained by the nature of the predictors used in this analysis. Save for age, Grade and Slum for the girls do not have many levels to allow for more 'continuous' predictors. With more variability within this predictors it is possible to achieve even greater accuracy. Two of these predictors contributed most to either being raped or not. The slum from where the girls comes from and the girl's age contributes most to the likelihood of getting raped. Studies with strong statistical approaches have over the years investigated just how significant or not some of these factors are across different contexts [15].

Employing logistics regression approaches, studies have demonstrated that age and sex are one of the main risk factors of rape. In a study conducted in England, multivariate logistic regression identified sex and age as one of the most significant demographic risk factors for rape [14]. Using such results it is possible to estimate the risk of sexual assault for different groups. From the demographic data collection from the 4 slums, the populations are more or less the same making it possible to generalize the findings of the study.

These results are consisted with others that have demonstrated that Age, level of education (grade) and socio-economic status as a predictor of sexual assault. From the results with respect to empowerment it is very clear who the perpetrators target. As demonstrated by

the Zimbabwe study (Gregson et al, 2011) a girls empowerment in terms of the level of education determines her likelihood of getting sexually assaulted. As demonstrated by this study younger girls, that is Form 2s are the most targeted ($p<0.001$). A possible explanation to this is the fact that these girls are just starting to get sexually active and are not yet fully empowered or educated enough to know the dangers that come with this situation. As they get more educated and experienced around their sexuality the rapes reduce as demonstrated by the low numbers of sexual assault among Form 3s and 4s. Education and empowerment plays a pivotal role in the reducing sexual assault.

Empowered girls have been able to prevail over sexual assault. In fact in a meta-analysis of all intervention against sexual assault carried out in 2014, (Ellsberg et at, 2014) [23], it was demonstrated that interventions that sought to empower women were more effective in preventing sexual assault compared to the rest. From the systematic review of reviews that was carried out , evidence for effective interventions was highly skewed towards high-income countries and all these involved aspects of economic empowerment. More than 80% of the rigorous evaluations were done in six high-income countries (Australia, Hong Kong Canada, the UK, New Zealand and the USA), comprising 6% of the world's population. The USA alone accounted for two thirds of all the intervention studies [29, 33, 34]. In this review empowerment interventions came out strongly as the most effective towards addressing sexual assault. From our analysis lower levels of education are associated with higher levels of sexual assault.

With respect to age, our analysis demonstrates that within the bracket of school going girls, perpetrators prefers older girls, from *figure 5*, it evident that older girls were the most

targeted (p<0.001). One possible explanation to could be factor of sexual desirability. Older girls look 'ready" for sex according to the perpetrators.

Girls are raped by individuals close to them. *Figure 6* shows the perpetrators as reported by the victims. Over 65% of the girls were raped by Individuals close to them. Even though a lot of blame and suspicion is usually meted on the strangers around rape, there is need to focus interventions on sexual assault by individuals close to the victim. This results are similar to those found in other demographics in Africa and the west, in a study conducted in the US, it was found out that 79% of survivors were victimized by a close friend, an acquaintance or family member [33, 35]. These individuals later suffered emotional distress, and increased issues and problems at school. One of the effective interventions against rape is usually creation of safe spaces for girls and boys both at home and out of home environment (Usually schools) [33, 34]. Working with the premise that they will be assaulted by people they know will go a long way in addressing sexual assault.

From the results is evident that all the slums contribute significantly to the incidences of rape. The economic status across these slums is similar as demonstrated in table 1. Nairobi informal settlements are similar in many aspects. However smaller slums seems to contribute more rapes. Dandora and Huruma which are much smaller slums compared to Mukuru and Kibera contributed more rapes. Kibera and Mukuru have over the years been a focus of many NGOs and civil societies trying to address insecurity menace, they have therefore seen reduced insecurity incidences [3, 27]. Rape naturally also dropped.

Our results on poverty being a significant risk factor corroborates those witnesses in other areas and that have been demonstrated in other studies. Women coming from poor backgrounds are more at risk of sexual assault as they go about their day to day tasks [27,

35]. As they walk home from work or school they face heightened levels of getting sexually assaulted. One possible explanation is that children born to parents residing in resource limited environments usually have lesser parental supervision as their parents go about trying to make ends meet. These children are therefore vulnerable to attacks both at home and at school. Also, children from these environments a lot of times are working themselves to help supplement their parents earning. They therefore put themselves at risk of sexual exploitation while at it [19, 26]. Poverty forces many girls and women into occupations that put them at risks of sexual assault. In recourse limited environments these women are forced into transactional sex that usually carry with it heightened levels of sexual assault [19].Poverty puts a lot of pressure on girls to do whatever it takes to get a meal, education or a job. Girls therefore find themselves in situations where they have to accept whatever is offered. Perpetrators take advantage of this kind of situations [29].

The results from this paper corroborates those from literature, adolescent girls are at a heightened risk of sexual assault. A girl's age is an important factor of rape risk. Figure 5 shows that from this group of girls 17 and 18 year olds were the most targeted group. Other studies have demonstrated 90% of adult rape victims are female and that 82% of all juvenile victims are female [32]. Literatures shows that Females aged between 16 and 19 are four times more likely than the rest of the population to be victims of sexual assault or rape. Women aged between 18-24 who are college students are 3 times more likely than women in general to experience sexual violence.

Conclusion

Predictive statistics such as LDA can use a combination of information on age (sexual desirability), grade (Empowerment and education) and the slum (Economic status) from where a girl comes from to categorize a girl into either the sexually assaulted or not groups with a reasonable accuracy. With accurate random data collection, it is possible to assess a girl's risk of being sexually assaulted. The biggest contributors to the likelihood of being raped for girls in Nairobi slums are the slum from where a girl comes from, the girl's age and the grade the girl is in at school. With respect to grade, girls in form 2 are the most vulnerable and more research should be done to look into this phenomenon. Girls are generally more likely to be assaulted by individuals close to them as opposed to strangers. Generally, sexual desirability, Empowerment/education and Economic status of a girl or woman can be used as predictors of sexual assault with respect to the likelihood of girls being raped or sexually assaulted.

Recommendations

Sexual assault amongst adolescent is rampant and there is need to understand the forces that drive this vice. Even though focus on rape prevention revolves around older women there is need to have more concerted effort towards understanding the risk factors for adolescents. With the young un-empowered girls as the most likely to be sexually group there is need to look and study more in-depth why age and poverty are a strong risk factor. What aspects of these two predictors results in victimization? Our study demonstrates the need for a more targeted rape campaign that will put into consideration the aspect of Age appropriateness and economic empowerment. Prevention campaigns should not generalize the needs for adolescents and older women, adolescents have their own unique needs in preventing violence. More studies and interventions are needed to look into the high prevalence sexual assault meted on girls just getting into sexually active age groups.

References

1. Federation of Women Lawyers in Kenya.Gender-based domestic violence in Kenya.2008.Availableat:http://fidakenya.org/wpcontent/uploads/2011/02/Domestic-Violence. pdf. Accessed November 26, 2012

2. Kenya National Bureau of Statistics (KNBS) and ICF Macro. Kenya demographic and health survey 2008–2009. Calverton, MD: KNBS and ICF Macro; 2010. Available at: www. measuredhs.com/pubs/pdf/FR229/FR229.pdf

3. Sinclair J, Sinclair L, Otieno E, Mulinge M, Kapphahn C, Golden NH. A self-defense program reduces the incidence of sexual assault in Kenyan adolescent girls. J Adolesc Health. 2013;53(3):374–380

4. Fonck K, Leye E, Kidula N, Ndinya-Achola J, Temmerman M. Increased risk of HIV in women experiencing physical partner violence in Nairobi, Kenya. AIDS Behav. 2005;9 (3):335–339

5. World Health Organization. Multi-Country Study on Women's Health and Domestic ViolenceAgainst Women: Summary Report of Initial Results on Prevalence, Health Outcomes and Women's Responses. Geneva, Switzerland: World Health Organization; 2005

6. Krishnan S, Dunbar MS, Minnis AM, et al. Poverty, gender inequities, and womens risk of human immunodeficiency virus/AIDS. Ann N Y Acad Sci 2008;1136:101e10.

7. Black MC, Basile KC, Breiding MJ, et al. The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 summary report. Atlanta (GA): National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2011.

8. Williamson N, Blum R. Motherhood in Childhood: Facing the Challenge of Adolescent Pregnancy. State of World Population. In: UNFPA, editor. 2013

9. Nairobi. CftSoA. Down the drain: counting the cost of teenage pregnancy and school drop out in Kenya. 2008.

10. Lloyd C, Young J. New Lessons: The Power of Educating Adolescent Girls. New York: The Population Council, 2009

11. Jewkes R, Sikweyiya Y, Morrell R, Dunkle K. Gender inequitable masculinity and sexual entitlement in rape perpetration South Africa: findings of a cross-sectional study. PloS one. 2011;6(12).

12. Whitaker DJ, Le B, Hanson RK, Baker CK, McMahon PM, Ryan G, et al. Risk factors for perpetration of child sexual abuse: A review and meta-analysis. Child Abuse and Neglect. 2008;32:529–48. doi:10.1016/j.chiabu.2007.08.005

13. Jewkes R, Fulu E, Roselli T, Garcia-Moreno C. Prevalence and risk factors for non-partner rape perpetration: findings from the UN Multi-country Cross-sectional Study on Men and Violence in Asia and the Pacific. The Lancet Global Health. 2013. Epub September 10, 2013 doi:http://dx.doi.org/10.1016/S2214-109X(13)70069-X e1

14. Brooker, C., & Tocque, K. (2016). Mental health risk factors in sexual assault: What should Sexual Assault Referral Centre staff be aware of? Journal of Forensic and Legal Medicine, 40, 28-33. doi:10.1016/j.jflm.2016.01.028

15. R, J., Y, S., K, D., & R, M. (2015). Relationship between single and multiple perpetrator rape perpetration in South Africa: A comparison of risk factors in a population-based sample. BMC Public Health, 15, 616. http://doi.org/10.1186/s12889-015-1889-9

16. Niehaus I. Towards a Dubious Liberation: Masculinity, Sexuality and Power in South African Lowveld Schools, 1953-1999" J South Afr Stud. 2000;26(3):387–407. doi: 10.1080/713683581

17. Wojcicki J. "She drank his money": survival sex and the problem of violence in taverns in Gauteng province, South Africa. Med Anthropol Q. 2002;16:1–28. doi: 10.1525/maq.2002.16.3.267

18. Shamu, S., Gevers, A., Mahlangu, B. P., Shai, P. N., Chirwa, E. D., & Jewkes, R. K. (2015). Prevalence and risk factors for intimate partner violence among Grade 8 learners in urban South Africa: Baseline analysis from the Skhokho Supporting Success cluster randomised controlled trial. International Health Int. Health. doi:10.1093/inthealth/ihv068

19. Gregson S, Zhuwau T, Ndlovu J, et al. Methods to reduce social desirability bias in sex surveys in low-development settings: Experience in Zimbabwe. Sex Transm Dis 2002;29:568e75.

20. S. Kim, A. Magnani, and S. P. Boyd. Robust Fisher Discriminant Analysis. In Y. Weiss and B. Sch¨olkopf and J. Platt, editor, Advances of the 18th Conference on

Neural Information Processing Systems (NIPS 2005), volume 18, pages 659 – 666, December 2005

21. Police. Kenya annual crime report for the year 2011. 2012. Available at: www.kenyapolice. go.ke/resources/CRIME%20REPORT%202011.pdf.

22. Brown DW, Riley L, Butchart A, Meddings DR, Kann L, Harvey AP. Exposure to physical and sexual violence and adverse health behaviours in African children: results from the Global School-Based Student Health Survey. Bull World Health Organ. 2009;87(6):447–455

23. Mary Ellsberg, Diana J Arango, Matthew Morton, Floriza Gennari, Sveinung Kiplesund, Manuel Contreras, Charlotte Watts Prevention of violence against women and girls: what does the evidence say? The Lancet , Volume 385 , Issue 9977 , 1555 – 1566

24. V. Diciunas, S. Raudys, Generalization error or randomized linear empirical error classifier,Informatica 11 (2000) 381–396.

25. N. Glick, Additive estimators for probabilities of correct classification, Pattern Recognition 10(1978) 211–222.

26. Namy S, Heilman B, Stich S, Edmeades J. Be a man, change the rules! Findings and lessons from seven years of CARE International Balkans' Young Men initiative. Washington, DC: International Center for Research on Women, 2014.

27. Ricardo C, Eads M, Barker G. Engaging boys and young men in the prevention of sexual violence. Pretoria: Oak Foundation, 2011.

28. Solotaroff J, Pande R. Violence against women and girls: lessons from south Asia. Washington, DC: World Bank, 2014

29. Abramsky T, Watts CH, García-Moreno C, et al. What factors are associated with recent intimate partner violence? Findings from the WHO Multi-Country Study on women's health and domestic violence. BMC Public Health 2011; 11: 109.

30. Heise L. Determinants of partner violence in low and middle-income countries: exploring variation in individual and population-level risk [PhD thesis]. London: London School of Hygiene & Tropical Medicine, 2012

31. Koenig MA, Ahmed S, Hossain MB, Khorshed Alam Mozumder AB. Women's status and domestic violence in rural Bangladesh: individual- and community-level effects. Demography 2003; 40: 269–88.

32. Grabe S. Promoting gender equality: the role of ideology, power, and control in the link between land ownership and violence in Nicaragua. Anal Soc Issues Public Policy 2010; 10: 146–70.

33. Solorzano I, Bank A, Pena R, Espinoza H, Ellsberg M, Pulerwitz J.Catalyzing personal and social change aroung gender, sexuality, HIV: impact evaluation of Puntos de Encuentro's communication strategy in Nicaragua. Washington, DC: Population Council, 2008.

34. Duflo E, Dupas P, Kremer M, Sinei S. Education and HIV/AIDS Prevention: Evidence from a Randomized Evaluation in Western Kenya. Background Paper to the 2007 World Development Report. World Bank Policy Research Working Paper 4024. Washington, DC: World Bank, 2006.

35. Bandiera O, Buehren N, Burgess R, Goldstein M, Gulesci S, Rasul I, et al. Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa. London, UK: 2014

36. Baiocchi M, Cheng J, Small D. Tutorial in biostatistics: Instrumental variable methods for causal inference. Statistics in Medicine. 2014:2297-340

37. Sarnquist C, Omondi B, Sinclair J, Gitau C, Paiva L, Mulinge M, et al. Rape Prevention Through Empowerment of Adolescent Girls. Pediatrics. 2014; epub ahead of print.

Appendix

Appendix 1: Data collection Questionnaire

QNO:	AREA:

UJAMAA AFRICA - NO MEANS NO WORLDWIDE
PRE- CLASS SURVEY

This survey is about your views on health related issues. The information you give will be used for research purposes only. Please do **NOT** write your name on this survey. The answers you give will be kept private. No one will know how you answer. Try to be as honest as you can.

Completing the survey is voluntary. Your grade or mark in this class will not be affected by whether or not you answer the questions. If you do not want to answer any question, you may leave the room.

DATE:_____ NAME OF
SCHOOL:_____
What Class are you in: Form 1 ☐   **Form 2** ☐   **Form 3** ☐   **Form 4** ☐

What Month were you born? Jan☐   **Feb** ☐   **Mar** ☐   **Apr** ☐**May**☐**Jun**☐

Jul☐**Aug**☐ **Sep**☐   **Oct**☐   **Nov**☐   **Dec** ☐

What year were you Born: 1993☐   **1994** ☐ **1995**☐   **1996** ☐   **1997**☐
1998 ☐

For these questions, please choose **the appropriate response**. Circle the letter before your answer e.g. (A) YES   B.   NO

1. Have you taken **NO Means NO SD classes** before today? **A.** YES   **B.** NO

2. What are the **four primary targets** to hit on an attacker? (Choose only 4)
    A. Arm   **B.** Eyes   **C.** Legs   **D.** Throat   **E.** Knee   **F.** Groin
    G. Head   **H.** Jaw

3. If a big man tries to assault me, I know how to defend myself: **A.** True   **B.** False
    C. Not Sure

4. If I am grabbed by an assailant I should focus on what, so as to get free?
    A. The part of my body that is being held   **B.** What is in my purse that I
    can use as a weapon

    C. The parts of my body that is still free to fight with. **D.** Whether he is drunk or high

5. The main goal of self defense is to:
    A. Fight the assailant   **B.** Get the police   **C.** Get away
    D. Win

6. What are the best methods for you to use if you are attacked by a rapist?
    A. Lie   **B.** Scream   **C.** Scratch the eyes   **D.** break the knee   **F.**
    all are correct

7. Which of these is the most important for you to use in an assault?
    A. Eyes   **B.** Spirit/will to live   **C.** Mind   **D.** Voice   **E.** Body   **F.**
    All are correct

8. If there are multiple assailants, and I need to fight I should target:
    A. The most Scared   **B.** The weakest assailant   **C.** The leader   **D.** The
    Strongest assailant

9. It is ok to use force and even injure my boyfriend if he is forcing me to have sex and
    won't listen to me.
    A. True   **B.** False   **C.** Not Sure

10. Have you **ever** been forced against your will to have sex (penetration of your vagina,
    anus or mouth with a man's penis or another object)? **A.** YES   **B.** NO

        a.   How many times?   **A.** 1 Time   **B.** 2 Times   **C.** 3 Times   **D.** 4 Times   **E.**
             5 Times   **F.** Never
        G. Other_____

b. Who forced you to have sex ? **A.** NEVER FORCED   **B.** NEIGHBOUR **C.** STEPFATHER/FATHER   **D.** BROTHER **E.** RELATIVE **F.** STRANGER **G.** TEACHER   **H.** PASTOR   I. GANGSTER **J.** POLICE   **K.** DOCTOR **L.** FRIEND **M.** BOYFRIEND   **L.** OTHER_____

c. Did you tell someone else about it? **A.** YES   **B.** NO   **C.** NEVER FORCED

d. If yes, whom did you tell? **A.** NEVER FORCED   **B.** NEIGHBOUR   **C.** RELATIVE   **D.** TEACHER   **E.** PASTOR **F.** POLICE **G.** DOCTOR **H.** FRIEND **I.** BOYFRIEND **J.** OTHER_____

11. In the past **12 months**, have you ever been forced against your will to have sex (penetration of your vagina, anus or mouth with a man's penis or another object)? **A.** YES **B.** NO

a. How many times?   **A.** 1 Time **B.** 2 Times **C.** 3 Times **D.** 4 Times **E.** 5 Times   **F.** Never
G. Other_____

e. Who forced you to have sex?   **A.** NEVER FORCED   **B.** NEIGHBOUR **C.** STEPFATHER/FATHER   **D.** BROTHER **E.** RELATIVE **F.** STRANGER **G.** TEACHER   **H.** PASTOR   **I.** GANGSTER   **J.** POLICE   **K.** DOCTOR **L.** FRIEND **M.** BOYFRIEND   **L.** OTHER_____

b. Did you tell someone else about it?   **A.** YES   **B.** NO   **C.** DOES NO APPLY TO ME

c. If yes, whom did you tell ?   **A.** NEIGHBOUR **B.** RELATIVE **C.** TEACHER **D.** PASTOR   **E.** POLICE **F.** DOCTOR **G.** FRIEND **H.** BOYFRIEND **I.** NEVER FORCED   **J.** OTHER_____

d. Why did you tell Him/her…………………………………………………….

12. What is the name of your upcountry district?_____

13. What was the name of your first primary school?_____

14. What is the name of the school where you did your K.C.P.E?-
_____

15. The first letter of your first (**Christian/Muslim**) name lies between?

1. A-G   **2.** H-N   **3.** O-T   **4.**U-Z

Appendix 2: Participants consent form

SCHOOL NAME: _____

NMNW's STANDARD CONSENT SCRIPT

This is read to the participants before the first, 3rd and 6th class (which is when the questionnaires are administered**.** Also before the refresher classes, which have questionnaires too.

Hi everyone, my name is……………… and I'm from Ujamaa Africa. Our goal is to teach you simple verbal and physical techniques to enable you to defend yourself against sexual assault. Your participation in this class is voluntary. We have been teaching this class for the past 5 years and all the girls who have taken it have really enjoyed the experience. We do exercises that are fun, you will learn many important physical skills that will be fun to learn and will certainly help you to protect yourself if you ever find yourself in danger. We are also conducting a research study on how effective these classes are. The purpose of the research is to see how much these classes help you stop harassment and sexual assault. We are also interested in assessing the general health of school children in your district. If you agree to participate, you will be contributing to the safety of thousands of women and girls who are most at risk of sexual violence. We will ask you to fill out a brief questionnaire at the end of some of the classes. The questionnaire takes about 20 minutes to complete and we will guide you through it. This questionnaire is completely

confidential. No one will ever know exactly who it was who answered which questionnaire, even us, the trainers. You don't put your name on it and we don't ask. You simply fold the questionnaire up when you're done and stuff it in that locked box at the front of the room. We take it to our offices and someone else empties it out. They or we will never know who it is that filled out which questionnaire. Participating in this research is voluntary and refusing or discontinuing participation will not lead to any penalty for you, from us or your teachers. If you have any questions you can ask any of the trainers here in class or we can even speak with you privately, you can also call this number for more clarifications. **0724 814 318.**

I ………………………………….. Do agree to participate on this research.

Name and Signature of Interviewer: Date:

_____   _____

Name of the student/Pupil   Date:

_____   _____