REGIONAL VARIATION AND FACTORS ASSOCIATED WITH FEMALE GENITAL MUTILATION AMONG WOMEN OF REPRODUCTIVE AGE IN ETHIOPIA: GENERALIZED ESTIMATING EQUATIONS AND MULTILEVEL ANALYSIS

BY: DESTA BEKELE AMENU

A Thesis submitted to School of Graduate studies Jimma University College of Natural Science Department of Statistics in the Partial Fulfilment of the Requirement for the Degree of Master of Science in Biostatistics

November, 2018

Jimma, Ethiopia

GEOGRAPHICAL VARIATION AND FACTORS ASSOCIATED WITH FEMALE GENITAL MUTILATION AMONG WOMEN OF REPRODUCTIVE AGE IN ETHIOPIA: GENERALIZED ESTIMATING EQUATIONS AND MULTILEVEL ANALYSIS.

BY

DESTA BEKELE AMENU

ADVISOR: GIRMA TAYE (PhD)

CO-ADVISOR: KIBREALEM SISAY (MSc)

November, 2018

Jimma, Ethiopia

DEPARTMENT OF STATISTICS, SCHOOL OF GRADUATE STUDIES JIMMA UNIVERSITY

As thesis research advisors, we hereby certify that we have read the thesis research prepared by DESTA BEKELE under our guidance, which is entitled "GEOGRAPHICAL VARIATION AND FACTORS ASSOCIATED WITH FEMALE GENITAL MUTILATION IN ETHIOPIA: Generalized Estimating Equation and Multilevel analysis", in its final format it is consistent and acceptable. Hence we recommend that the thesis accepted as it fulfils the university and department style requirements for the degree of Master of Science in Biostatistics.

Dr. Girma Taye (PhD)	____________________	______________
Main advisor	Signature	Date
Kibrealem Sisay (MSc.)	______________________	___________________
Co-advisor	Signature	Date

As the members of the board of examiners of MSc. thesis open defence examination of DESTA BEKELE AMENU, we certify that we have read and evaluated the thesis and examined the candidate. We recommend that the thesis be accepted as it fulfils the requirements for the degree of Master of Science in Biostatistics.

____	_________________	__________________
Name of chairman	Signature	Date
_______________________________________	___________________	_______________
Name of Main Advisor	Signature	Date
____	__________________	________________
Name of Co-advisor	Signature	Date
____	________________	_______________
Name of internal Examiner	Signature	Date
____	________________	_______________
Name of External Examiner	Signature	Date

DECLARATION

I declare that, this thesis is a result of my genuine work and all sources of materials used for writing it have been duly acknowledged. I have submitted this thesis to Jimma University in partial fulfilment for the Degree of Master of Science in Biostatistics. And also, I solemnly declare that I have not so far submitted this thesis to any other institution anywhere for that award of any academic degree, diploma or certificate.

Desta Bekele

Date: ________________

Signature: ____________

Jimma University, Jimma

ACKNOWLEDGMENT

First and foremost I would like to thank the almighty God for his support and strength throughout the process of carrying out this study and for seeing me through all.

Next, I would like to thank my thesis advisors Dr. Girma Taye and Mr. Kibrealem Sisay for their invaluable comment and suggestion.

And also, I express my heavy gratitude to my signora Derartu Marga for providing me with unfailing support and continuous encouragement throughout my education and through the process of researching and writing this thesis. This accomplishment would not have been possible without her. Thank you.

Finally, I am also very grateful to my friend Busha Giduma whose presence and support have always been important to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYM

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CIC	Correlation Information Criterion
CSA	Central Statistical Agency
DHS	Demographic Health Survey
EA	Enumerate Area
EDHS	Ethiopia Demographic Health Survey
FGC	Female Genital Cutting
FGM	Female Genital Mutilation
GEE	Generalized Estimation Equation
GLM	Generalized Linear Model
GLMM	Generalized Linear Mixed Model
ICC	Intra Class Correlation
IPPF	International Planned Parenthood Federation
MDG	Millennium Development Goals
MML	Marginal Maximum Likelihood
PHC	Population and House Census
STI	Sexual Transmitted Infection
UNICEF	United Nation International Children Fund
UNFPA	United Nation Female Population Association
WHO	World Health Organization
Fij	Fully Iterated Jack-knife Standard Error

ABSTRACT

Background: Female Genital Mutilation (FGM) is one of harmful traditional practice in developing countries including Ethiopia. This practice causes significant and irreversible damage to the physical, psychological and sexual health of many women and girls and is one of the most devastating human rights violations. Generalized estimating equations are an extension of GLMs to accommodate correlated data. The focus of the GEE is on estimating the average response over the population rather than the regression parameters that would enable prediction of the effect of changing one or more independent variables on a given individual. Multilevel analysis is a methodology for the analysis of data manifesting complex variability, with a focus on the nested source of variability.

Objective: The objectives of this study are to identify factors that influence female genital mutilation, the regional variability of female genital mutilation in Ethiopia, and modelling female genital mutilation using generalized estimating equations and multilevel models.

Methodology: Data of Ethiopian demographic and health surveys (EDHS) of 2016 was used in this research. It includes nationally representative of 16,583 ever married women aged 15-49. Hot-deck multiple imputations were used to handle missing in data and improve the reliability of the inference. Generalized Estimating Equations and Multilevel Logistic Regression were carried out to analyse covariates related to FGM among women and daughters' included in the study with statistical package R.

Result and Conclusion: The results obtained from the generalized estimating equation and multilevel logistic regression showed that age, type of residence, religion and education level significantly associated with female genital mutilation among women and there is a variation of female genital mutilation across region. In addition to that age of mother and circumcision status of mother's significantly related with female genital mutilation among daughters and there is variation of female genital mutilation across the region. We compared two model (GEE and multilevel) to identify the model well describe the association of explanatory and response variables. Using standard error corresponding parameters, the multilevel analysis is used for further discussion. It was also found that a random intercept model was the best description of the data set among multilevel models.

Key Words: Female Genital Mutilation/Cutting; Generalized Estimating Equations (GEE); Working correlation; Multilevel

CHAPTER ONE

1. INTRODUCTION

1.1. Background of the study

Female genital mutilation (sometimes called female genital cutting) is defined by the WHO as referring to all procedures involving partial or total removal of the external female genital organs for non-medical reasons (for cultural, religious or other non–remedy reasons). FGM is a form of gender-based violence and has been recognised as a harmful practice and a violation of the human rights of girls and women. More than 200 million girls and women globally are estimated to have undergone FGM/C. Over 3 million girls are estimated to be at risk of undergoing FGM annually[1,2].

During the past three decades, several international and national humanitarian and medical organizations have drawn worldwide attention to the physical harms associated with FGM. The World Health Organization and the International Federation of Gynaecology and Obstetrics have opposed FGM as a medically unneeded practice with serious, potentially ominous complications[3]. The American College of Obstetricians and Gynaecologists and the College of Physicians and Surgeons of Ontario, Canada, also opposed FGM and advised their members not to perform these procedures. In 1995 the Council on Scientific Affairs of the American Medical Association recommended that all physicians in the United States strongly denounce all medically unnecessary procedures to alter female genitalia, as well as promote culturally sensitive education about the physical consequences of FGM[4,5].

Practices of FGM/C have been found throughout history in many cultures, but there is no classic evidence documenting when or why this practice began. The origin is thought to precede the rise of Christianity and Islam. Egyptian mummies have been described displaying characteristics of FGM/C, and it is thought that FGM/C may have been a sign of distinction amongst the ruling class[6].

UNFPA conceived the idea of bringing together experts and practitioners to a global forum on female genital mutilation/cutting, guided by its commitment to the abandonment of FGM/C, which has created suffering and pain among millions of women through generations. Indeed, this practice causes significant and irreversible damage to the physical, psychological and sexual health of many women and girls and is one of the most devastating human rights violations that are hidden from view[7]. FGM/C violates the human rights of infants, adolescent girls and women who are incapable of giving informed consent due to age or coercion. Recent studies indicate the profound effects of FGM/C on maternal morbidity and mortality and, possibly, on increased infant mortality[8]. The fear of stigmatisation within the host society cultural norms in regards to the integrity and rights of girls' and women's bodies as well as the open perceptions that FGC is a varied practice that requires criminalisation has been documented as inadvertently limit women and men from accessing needed quality health services[9].

In Ethiopia, the prevalence of FGM among women was 79.9% in 2000 and 74.3% in 2005. On the EDHS report of 2000, the prevalence rate of FGM among rural and urban populations was relatively similar. However, in the EDHS 2005, the prevalence of FGM among the urban and rural population was 68.5% and 75.5% respectively. The prevalence rate of FGM also showed regional variation. The highest prevalence was documented in Somali region both in 2000 and 2005 EDHS[10].

Generalized Estimating Equations (GEE) and Multilevel Logistic Regression were used to model the female genital mutilation/cutting practices. Multilevel models (also known as hierarchical linear models, nested data models, mixed models) are statistical models of parameters that vary at more than one level. In this research, a model of women that contains circumcision status for individual women and their daughters as well as circumcision status for the region within which the women are grouped. This study is intended to investigate the factors which influence female genital mutilation by considering regional heterogeneity as a random effect in the multi-level analysis.

Generalized estimating equations were introduced by Liang and Zeger (1986) as an extension of generalized linear models (GLM) to analyze discrete and correlated or data. Its strength is that it models a known function of the marginal expectation of the dependent variable as a linear function of explanatory variables. GEE estimates by marginal maximum likelihood (MML) the factors related to FGM/C which is one of the biggest social problems that is affecting the majority of women and young girls.

1.2. Statement of the problem

Female genital mutilation/cutting is one of the most dangerous practices that cause injury and death among those who undergo the procedure. FGC can induce a range of health problems, both short-term and long-term. Short-term health problems like bleeding, infections, pain, and trauma. Long-term health problems like problems with gynaecological health, increased the risk of sexually transmitted infections (STIs) including HIV, Psychological and emotional stress, problems getting pregnant and problems during pregnancy[11].

According to the Demographic Health Survey of Ethiopia, the estimated prevalence of FGM in girls and women (15-49 years) is 74.3% (DHS, 2005). The demographic health survey of 2011 has not been collected FGM data. 23.8 million Women and girls in Ethiopia have undergone FGM, and this is the second highest rate in Africa, next to Egypt. FGM is widespread across Ethiopia and is carried out in the majority of regions and ethnic groups[10].

And due to the multi-ethnic and multi-cultural nature of the society, the prevalence of FGM across regions of Ethiopia and factors related with FGM may cause variation at regional levels. Thus, this study tries to address the regional variation of FGM and explore the risk factors associated with genital mutilation (cutting) of daughters and women under reproductive ages taking into consideration socio-economic and environmental factors such as mother's educational level, father's level of education, place of residence, religion, media exposure, age of women, economic status of the household based on the 2016 Ethiopia Demographic and Health Survey data.

Generally, this study answers the following basic research questions:

- What is the prevalence of female genital cutting over all the country?
- Which factors that have an impact on genital mutilation practice?
- In which region female genital mutilation practice highly distributed?
- Which statistical model is appropriate for modelling determining factors and regional variation regarding female genital cutting?

1.3. Objective

1.3.1. General Objective

The main objectives of this study were to identify factors that are associated with female genital circumcision and geographic variability of female genital mutilation in Ethiopia.

1.3.2. Specific Objectives

1. To estimate the prevalence of female genital mutilation/cutting in Ethiopia.
2. To identify factors associated with female genital mutilation practices.
3. To assess regional variation regarding female genital mutilation/cutting in Ethiopia.
4. To compare results from generalized estimating equations and multilevel logistic regression model in determining factors of female genital mutilation.

1.4. The significance of the study

The findings of this study were focused on:

1. To model and give emphases on the factors that have a strong association with female genital mutilation (cutting) so that policy directives act on accordingly.
2. To show the regional variation of FGC/M practiced and which give directions to stakeholders.
3. Most of researcher and statistical packages use listwise deletion method of handling missing data rather than imputation. This study promotes hot-deck multiple imputations to the researcher to reduce bias and improve reliability using R statistical software.

CHAPTER TWO

2. LITERATURE REVIEW

2.1. Literature related to the variables used in the study

Female genital mutilation (FGM) is a comprehensive term that includes all procedures that involve partial or total removal of the external female genitalia, or another injury to the female genital organs, for non-medical reasons. Female genital mutilation violates a number of well-established human rights principles, norms and standards, including the principles of equality and non-discrimination on the basis of sex, gender, the right to bodily integrity, the right to life (because the procedure can result in death), and the right to the highest attainable standard of physical and mental health. Female genital cutting or female genital mutilation (FGC/M) is terms used deliberately by some activists specifically to encourage practising communities to desert the practice. The term FGM is used in this Statement to emphasize the serious physical, emotional and psychological consequences associated with the procedure[3,7].

Female genital mutilation/cutting includes “a range of practices involving the complete or partial removal or damage of the external genitalia for non-medical reasons”. This procedure may involve the use of unsterilized, makeshift or rudimentary tools. Female genital mutilation/cutting (FGM/C) is a global concern. Not only is it practiced among communities in Africa and the Middle East, but also in immigrant communities throughout the world. Moreover, recent data reveal that it occurs on a much larger scale than previously thought. It continues to be one of the most persistent, pervasive and silently endured human rights violations. FGM is a form of gender-based violence and has been recognized as a harmful practice and a violation of the human rights of girls and women. More than 200 million girls and women worldwide are living with the effects of FGM[2].

According to thesis of six sub Saharan African countries by Kerubo, Karhu Rose, female genital mutilation-effects on women and young girls, Ethiopia out lawed female genital mutilation in 2004, but still the practice is deeply rooted and nearly universal in the country. In 2005, a government health survey of the Ethiopia country found out that 74 percent of girls and women had undergone the ritual cutting[12]. Over 125 million girls and women alive today have had FGM in the 30 African countries and 3 million girls are estimated to be at risk of undergoing FGM annually[2].

Prevalence

In 28 countries in Africa and the Middle East for which data are available, national prevalence among women aged 15 years and older ranges from 0.6% (Uganda, 2006) to 97.9% (Somalia, 2006). There are some regional patterns in FGM prevalence. According to Demographic Health Surveys done during 1989–2002, within north-eastern Africa (Egypt, Eritrea, Ethiopia, and Northern Sudan), prevalence was estimated at 80–97%, while in eastern Africa (Kenya and the United Republic of Tanzania) it was estimated to be 18–38%. However, prevalence can vary strikingly between different ethnic groups within a single country. FGM has been documented in several countries outside Africa but national prevalence data are not available[10].

In Ethiopia, 79.9% in 2000 and 74.3% in 2005 women have been circumcised. Among these women, the most common type of FGM/C involves the cutting and removal of flesh (73%). FGM/C is more common among women from rural areas (75.5%) than urban areas (68.5%). Regionally, FGM/C is least common in Tigray (29.3%) and Gambella (27.1%) and more common in Afar (91.6%) and Somali (97.3%). FGM/C has declined from 80% of women in 2000 to 74% in 2005[13].

Social consequences

The FGM practice is performed in response to strong social conventions and supported by key social norms; thus failure to conform often results in harassment and exclusion from important community events and support networks, as well as discrimination by peers. Unless there is a joint agreement within a larger group, individuals and families are likely to consider the social risks to be greater than the physical and mental health risks to girls of FGM. Even legal restrictions against FGM may be seen as less important than the restrictions that can be imposed by the community for non-compliance with the practice[10].

Health consequences

FGM has no health benefits. It involves removing and damaging healthy and normal female genital tissue, and interferes with the natural functions of girls' and women's bodies. Traditional practices use a variety of tools to perform FGM, including razor blades and knives, and do not usually use anaesthetic. An estimated 18% of all FGM is done by health-care providers, who use surgical scissors and anaesthetic. All forms of FGM can cause immediate bleeding and pain and are associated with risk of infection; the risk of both immediate and long-term complications increases with the extent of the cutting. Research into the health effects of FGM has progressed in recent years. A WHO led a study of more than 28,000 pregnant women in six African countries found that those who had undergone FGM had a significantly higher risk of childbirth complications, such as caesarean section and postnatal bleed, than those without FGM. In addition, the death rate for babies during and immediately after birth was higher for mothers with FGM than those without. The risks of both birth complications and neonatal death increased relative to the severity of the type of FGM. Sexual problems are also more common among women who have undergone FGM. They are 1.5 times more likely to experience pain during sexual intercourse, have the significantly less sexual satisfaction and are twice as likely to report a lack of sexual desire[5].

Women who had undergone FGM/C were significantly more likely to have Caesarean sections, risks for extensive bleeding; longer hospital stays after delivery, prolonged labour, and the need for episiotomies, death, and the resuscitation of the infant and low birth weight. Thus, the practice of FGM/C contributed to the high rates of maternal death reported in Africa as well as infant deaths. FGM/C is, therefore, a global concern[8].

Cultural Requirements

Many justifications are given for FGM; the reasons are complex, and vary by country, region and ethnicity, even within communities. It is entrenched in social, economic, cultural and political structures and understood as a social convention that is often accepted without question. Some of the social justifications include the preservation of virginity and ensuring fidelity, as well as a rite of passage to womanhood in some contexts. The practice can therefore be construed as an important part of the cultural identity of girls and women[8].

The traditions and beliefs have continued to be stronger in most African regions and many have become strict followers of their cultures, which is one of the reasons why the practice is still practiced. According to Our Selves, Our Daughters Community-Based Education and Engagement Addressing Female Genital Cutting (FGC) with Refugee and Immigrant African Women in Winnipeg –2010-2011 of Final Activity and Evaluation Report (April 2011) results, in multi-ethnic societies across the western world, professionals in the field of health and social services are faced with an increasing number of women, men, and families originating from countries where practices such as FGC are common. International trends of migration contribute to the growing controversy regarding traditional practices as they meet up with host society's cross-cultural imperatives in the health care system. As such, health care professionals are required to deal with the ethical complexity of navigating through their own personal identity and culture, most often in opposition with the identity and cultural processes of the women and men migrants they are meant to serve while bound by their legal and professional guidelines[14].

Religious Requirements

According to Global Consultation on Female Genital Mutilation/Cutting of Technical Report (2008), religious justifications across Christian, Jewish, Muslim and some indigenous African groups are often invoked for the practice, although none of the Holy Scriptures in any of these religions dictates female genital mutilation. Understanding these cultural and societal beliefs is a critical element in any work that aims to eliminate the harmful practice[8].

FGM is practiced across religions including Christians, Jews, Animists, and Muslims. Within Muslim communities, religion is a commonly cited reason for FGM/C. Female circumcision is not mentioned in the Koran. However, a much-disputed reference to it may exist in the Sunna, which is a collection of the words and actions of the Prophet Mohammed[6].

The dissertation of Karhu Rose Kerubo (Järvepää Autumn 2010) on Female Genital Mutilation and its effects on women and young was concluded that various religions such as Islam or Christianity are not in favour of the continuation of the practice. From the interviews, based on the analysis and information from the interviews, the women are told the practice has religious justification and therefore they had to do it if they wanted to be Muslim women[12].

The articles of Geographic variation and factors associated with female genital mutilation among reproductive-age women in Ethiopia: based on a national population-based survey describes that, Muslim women had 3 times higher odds of having experienced FGM as compared to Orthodox women ($p<0.0001$). Women in the higher age categories had higher experiencing FGM as compared to young women (15–19 years). Daughters of richer women, Muslim mothers and mothers aged ≥25 years had higher of having experienced FGM[13].

As Setegn, T, Lakew, Y and Deribe, K (2016) on Geographic variation and factors associated with female genital mutilation among reproductive age women in Ethiopia: based on a national population based survey results, daughters of Protestant women had lower odds of having experienced FGM. Higher levels of maternal education were associated with 80% lower odds of FGM experience for daughters[13].

Education Gap

Education, especially of women, can play an important role in safeguarding the human rights of both women themselves, and those of their children. Overall, daughters of mothers who are more highly educated are less likely to have undergone FGM/C than daughters of mothers with little or no education[5,1].

Female genital mutilation is recognized both internationally and locally to be an enduring tradition which is difficult to overcome because it violates the rights of women and young girls. The fact that those letting their children undergo the procedure they do not know if they are violating the rights of the children because of the high level of illiteracy involved[7].

Geographic Variation

The article by Kidanu Gebremariam, Demeke Assefa, and Fitsum Weldegebreal(2009) depicted that the prevalence of female circumcision was a significant predictor among rural respondents when compared to their counterparts in urban respondents. It was four times higher among rural respondents than those who lived in urban areas (AOR =4; 95% CI =2, 6.8). This is in line with the report of EDHS 2005. This may be because of the tight tradition and religious association and lose legal concern for the practice in the rural areas and may become a convenient place for the practice; additionally, rural women are mostly influenced by traditional practices. This may be as most anti-FGC interventions, health promotion through media, and legally punishable action might not be addressed in rural areas as in urban areas[15].

According to the research conducted by Roman Asefa (2011) on factors affecting the practice of female genital mutilation of Ethiopian women results, the odds of daughters being circumcised has decreased by a factor of 0.028 for mothers living in Tigray compared to those in Dire Dawa and the odds of daughters being circumcised has decreased by a factor of 0.357 for mothers living in Afar compared to Dire Dawa. The odds of daughters being circumcised has decreased by a factor of 0.734 for mothers living in urban compared to those in the rural [16].

Wealthy Index Factor

The journal written by Setegn, T, Lakew, Y and Deribe, K (2016) on Geographic variation and factors associated with female genital mutilation among reproductive age women in Ethiopia: based on a national population based survey describes, Women in the richest and richer wealth index categories had higher odds of having experienced FGM as compared to women in the poorest category. Daughters from richer women depicted 40% higher odds of having experienced FGM as compared to daughters of the poorest women. Daughters of women in the higher age categories showed higher odds of having experienced FGM as compared with daughters of young women aged 15–19 years[13].

Age

As Setegn, T, Lakew, Y and Deribe, K (2016) on Geographic variation and factors associated with female genital mutilation among reproductive age women in Ethiopia: based on a national population based survey results, daughters of women in the higher age categories showed higher odds of having experienced FGM as compared with daughters of young women aged 15–19 years[13].

EDHS, 2016: In Ethiopia, FGM/C is performed throughout childhood. Women are most likely to report circumcision occurred before age 5 (49%), while 22% are circumcised between age 5-9, 18% age 10-14, and 6% age 15 or older[11].

2.2. The model used in the study

Generalized Estimating Equation

Generalized Estimating Equation (GEE) introduced by Liang and Zeger. They are an extension of Generalize Linear Model (GLM) to clustered and longitudinal analysis using quasi-likelihood. GEE treats covariance structure as a nuisance and GEE is not concerned about the variance of each data. Besides, GEE has often used as a general and computationally convenient method[17].

GEE is a marginal (or population-averaged) as opposed to a cluster-specific (or subject-specific, conditional) method. Hence, GEEs model the marginal expectations of the outcome and do not specify the joint distribution of a group’s observations. The method is most sensitive when the chief interest lies in the regression equation for the marginal expectations and not in the intra-cluster correlation structure. Longitudinal research often aims at describing the marginal expectations of the outcome as a function of the predictors. For example, when comparing a group of students subject to a novel teaching technique with a control group on a scholastic performance indicator over an extended period of time, the focus rests not on individuals' odds ratios but on the average odds ratios of the two groups. Population averaged methods model the "average response over the subpopulation that shares a common value of" the predictors as a function of such predictors (Diggle et al., 2002). Hence, they are appropriate when we have reasons to believe that few values of the predictive variables are shared by many observations[18].

The focus of the GEE is on estimating the average response over the population rather than the regression parameters that would enable prediction of the effect of changing one or more covariates on a given individual. GEEs are usually used in conjunction with Huber-White standard error estimates, also known as "robust standard error" or "sandwich variance" estimates. In the case of a linear model with a working independence variance structure, these are known as "heteroscedasticity consistent standard error" estimators. Indeed, the GEE unified several independent formulations of these standard error estimators in a general framework. The most popular form of inference on GEE regression parameters is the Wald test using naive or robust standard errors, though the score test is also valid and preferable when it is difficult to obtain estimates of information under the alternative hypothesis. The likelihood ratio test is not valid in this setting because the estimating equations are not necessarily likelihood equations. GEE takes into account the dependency of observations by specifying a "working correlation structure"[19,20].

Generalized Linear Models and multilevel logistic regression

To model types of responses such as binary and count data, we can use the generalized linear mixed-effects model. GLM extends linear regression for a continuous response to models for other types of response such as binary and categorical outcomes. Examples of GLMs include linear regression, logistic regression for binary outcomes, and log-linear regression for count data[21].

In multilevel generalized linear models, the multilevel structure appears in the linear regression equation of the generalized linear model. The multilevel regression model is more complicated than the standard single-level multiple regression models. One difference is the number of parameters, which is much larger in the multilevel model. This poses problems when models are fitted have many parameters, and in model exploration[22].

Multilevel analysis is a methodology for the analysis of data manifesting complex variability, with a focus on the nested source of variability. The best approach to the analysis of multilevel data is an approach that represents within-group as well as between group relation within a single level analysis, where “group” refers to the units at the higher levels of the nesting hierarchy. Probability models are used to represent the within-group and between-group variability. In another word, we conceive variation within groups and variation between groups as random variability[23,24].

Multilevel regression models are essentially a multilevel version of the familiar multiple regression model. Using dummy coding for categorical variables, it can be used to analysis of variance (ANOVA)-type of models as well as the more usual multiple regression models. Since the multilevel regression model is an extension of the classical multiple regression model, it too can be used in a wide variety of research problems. It has been used extensively in educational research[25,26].

CHAPTER THREE

3. MATERIALS AND METHODOLOGY

3.1. Sources of Data

The data used in this study is from the Ethiopia Demographic and Health Survey (EDHS) conducted in 2016, downloaded from the Measure DHS website (www.dhs.program.com) in SPSS format. The Ethiopia Demographic and Health Survey were implemented by the Central Statistical Agency (CSA) from January 18, 2016, to June 27, 2016. This survey is a nationally representative survey of 16,583 ever-married women aged 15-49 from 18,008 households were selected[27,11].

Administratively, Ethiopia is divided into nine geographical regions and two administrative cities. The sample for the 2016 EDHS was designed to provide estimates of key indicators for the country as a whole, for urban and rural areas separately, and for each of the nine regions and the two administrative cities.

The 2016 EDHS sample was stratified and selected in two stages. Each region was stratified into urban and rural areas, giving 21 sampling strata. Samples of EAs were selected independently in each stratum in two stages. Implicit stratification and proportional allocation were achieved at each of the lower administrative levels by sorting the sampling frame within each sampling stratum before sample selection, according to administrative units in different levels, and by using a probability proportional to size selection at the first stage of sampling.

In the first stage, a total of 645 EAs (202 EAs in urban areas and 443 EAs in rural areas) were selected with probability proportional to the EA size (based on the 2007 PHC) and with independent selection in each sampling stratum. A household listing operation was carried out in all the selected EAs from September to December 2015. The resulting lists of households served as a sampling frame for the selection of households in the second stage. In the second stage, a complete listing of households was carried out in each selected cluster. 28 households from each cluster were then systematically selected, and all women of reproductive age in the households are interviewed[27].

Thus, this study analyses responses from each of 15,683 women of age 15-49 interviewed in 2016, of practicing female genital mutilation. After understanding the detail datasets, all potential social determinants and female genital mutilation indicator variables were extracted from the Ethiopia demographic and health survey datasets.

Study Population

The study population in this study is all women within the reproductive age group (15-49) years who were living in Ethiopia during the study of the 2016 EDHS.

3.2. Variables Included in the Study

3.2.1. Dependent Variable

The dependent variables Y_{1i} (Circumcisions status of women among reproductive age) and Y_{2k} (Circumcisions status of daughters) are dichotomous random variables with "Circumcised" (coded as 1) and "not circumcised" (coded as 0). Thus, Y_{1i} and Y_{2i} takes on values, Y_{1i}= 0 and 1 where i denotes the individual mother and Y_{2k}= 0 and 1 where k denotes daughters' of i^{th} mother

$$\text{Women Circumcised status } (\mathbf{Y_{1i}}) = \begin{cases} 1 & if\ ith, women\ \ circumcised \\ 0 & Otherwise \end{cases}$$

$$\text{Daughters Circumcised status } (\mathbf{Y_{2K}}) = \begin{cases} 1 & if\ kth, daughters\ \ circumcised \\ 0 & Otherwise \end{cases}$$

3.2.2. Independent Variables

This study included the most important expected factors of female genital circumcision from various literature reviews and their theoretical justification from the source data.

Factors that influence female circumcision included in the study were: place of residence, religion, mother level of education, father level of education, region, mother age, wealth index, media exposure and region which included in Ethiopia Demographic and Health Survey of 2016[13,15 ,27].

No	Variable	Description and category
1	place of residence	1=urban 2=rural
2	Region	1=Tigray 2=Afar 3=Amhara 4=Oromiya 5=Somali 6=Ben-Gumuz 7=SNNP 8=Gambela 9=Hareri 10=Addis Ababa 11=Dire Dawa
3	Religion	1=Coptic orthodox 2=catholic 3=protestant 4=Muslim 5=traditional 6=others
4	mothers' level of education	0=no education 1=primary 2=secondary 3=higher
5	wealth index	1= poorest 2= poorer 3= middle 4=richer 5=richest
6	mother's Age	1=15-19 2=20-24 3=25-29 4=30-34

		5=35-39 6=40-44 7=45-49
7	Media exposure	0= Not exposed 1= exposed
8	Women circumcision status	0 = not circumcised 1 = circumcised
9	Fathers' level of education	0=no education 1=primary 2=secondary 3=higher 8=don't know

Table 3. 1 The independent variable of the study

3.3. Missing data

Missing data occur in survey research because an element in the target population is not included on the survey's sampling frame (non-coverage), because a sampled element does not participate in the survey (total non-response) and because a responding sampled element fails to provide acceptable responses to one or more of the survey items (item non response). Missing data commonly occur in demographic and health survey and are defined as no data values are stored for a variable or variables. Frequently, missing data can potentially lead to two serious problems in statistical practice:

i) Reducing the overall statistical power and having biases in the estimates.
ii) Missing data is a problem since nearly all standard statistical methods presume complete information for all variables included in the analysis.

The amount of bias potentially introduced by missing data depends on the type of missing mechanism.

Missing completely at random (MCAR): By stating that data are MCAR, we assume that the missing values are not systematically different from the values we did observe.

Missing at random (MAR): When data are MAR, the missing values are systematically different from the observed values, but the systematic differences are fully accounted for by measured covariates. In this situation, we can use what we know about partial cases to compensate for bias due to missing data.

Not missing at random (NMAR) or informative missing data: Concerns about NMAR data may be raised when missing values are thought to systematically differ from observed values. This can happen if (1) the missing value itself influences the probability of missingness or (2) some unmeasured quantity predicts both the value of the missing variable and the probability of missingness.

3.3.1. Handling Missing Data

Several methods and strategies are available to handle missing data. From them, listwise-deletion (delete cases), imputation, KNN-imputation, regression and other. Handling missing data has three advantages. First, the method yield unbiased estimates of a variety of different parameters. Second, the method include a way to assess the uncertainty about the parameter estimates, and third, the method should have good statistical power.

3.3.2. Multiple Hot-deck Imputation

Hot-deck multiple imputations was proposed by Rubbin (1987) and has been used by the census bureau to complete the public-use database. In multiple hot-deck imputations, several observed values of the variable with missing observations are drawn conditional on the rest of the data and are used to impute each missing value. The advantage of this class of methods over multiple imputations is that the imputed values are actually drawn from the observed data.

In hot deck imputation, missing values are imputed by copying values from similar records in the same dataset. Or, in notation:

$$\hat{x}_j = x_j,$$

Where x_j is taken from the observed values. Hot-deck imputation can be applied to numerical as well as categorical data but is only viable when enough donor records are available. The main question in hot-deck imputation is how to choose the replacement value x_j from the observed values. There are four well-known flavors. But we deal with random hot-deck imputation.

In random hot-deck imputation, a value is chosen randomly and uniformly from the same data set. When meaningful, random hot-deck methods can be applied per stratum. In this study, we apply random hot-deck imputation of circumcised and non-circumcised respondents separately[28]. Random hot-deck on a single vector can be applied with the impute function of the Hmisc package in R.

The most common means of dealing with missing data is a listwise-deletion, which is when all cases with a missing value are deleted. If the data are missing completely at random, then listwise deletion does not add any bias, but it does decrease the power of the analysis by decreasing the effective sample size.

Cold-deck imputation, by contrast, selects donors from another dataset. Due to advances in computer power, more sophisticated methods of imputation have generally replaced the original random and sorted hot-deck imputation techniques.

Another imputation technique involves replacing any missing value with the mean of that variable for all other cases, which has the benefit of not changing the sample mean for that variable. Mean imputation has some attractive properties for univariate analysis but becomes problematic for multivariate analysis. And our variable that need imputation is categorical. Thus, it is not appropriate to use in this study.

Hot-deck imputation involves replacing missing values of one or more variables for a non-respondent (called the recipient) with observed values from a respondent (the donor) that is similar to the non-respondent with respect to characteristics observed by both cases. It avoids the issue of cross-user inconsistency that can occur when analysts use their own missing-data adjustments. The hot-deck method does not rely on model fitting for the variable to be imputed and thus is potentially less sensitive to model misspecification than an imputation method based on a parametric model. In contrast to other listed above, hot-deck multiple imputation is the best to use in this study.

3.4. Method of Data Analysis

This study was exploring the data of EDHS conducted in 2016 related to female genital mutilation. This data was visualized in graph and table described at a different level as the structure of the EDHS data. The R software used to describe, impute missing and analyze data.

To assess the factors (demographic, socio-economic and social) related to female genital mutilation/cutting among the reproductive women, the covariates were put in the Generalized

Estimation Equations (GEE) and multilevel logistic Regression Model. And using 5% level of significance we assess their individual effect on FGM.

Marginal models and generalized linear mixed models are among the widely used models to study clustered or repeated data. GLMM is an extension to the generalized linear model (GLM) in which the linear predictor contains random effects in addition to the usual fixed effect.

Marginal effect models are models in which responses are averaged over all responses. The primary objective of the marginal model is to analyse the population-averaged effects of the given factors in the study on the response variable of interest. This means that the independent variables are directly related the marginal expectations[29]. In this study, the well-known GEE and GLMM were applied.

This study reviews two modelling strategies for binary data and their implementations in R. Generalized Estimation Equations and Multilevel Logistic Regression with their assumptions on variance.

Generalized Linear Model (GLM)

Generalized linear models (GLMs) extend ordinary regression models to encompass non-normal response distributions and modelling functions of the mean[19].
Let Y_i be the response and X_i explanatory variables or covariates for unit i, and define the conditional expectation of the response given the covariates μ_i, i.e. $\mu_i \cong E[Y_i|X_i]$.
Generalized linear models are specified as

$$g(\mu_i) = X_i^T\beta = \vartheta_i, \qquad 3.1$$

Where the linear combination $\vartheta_i = \beta_0 + \beta_1 X_{i1} + \ldots = X_i^T\beta$ is called "linear predictor" and β's are fixed effects. g is a "link function" linking the expected response μ_i, to linear predictor ϑ_i. The specification is competed by choosing a conditional distribution for the responses y_i given the covariates X_i, f(Y_i/X_i), from the exponential family distributions[29].

GLM assumes that the response variables are independent. In clustered data, like the EDHS data, observations will be taken from all women of reproductive age in the selected households, and female genital practice in a region may be correlated. Correlated data requires proper analysis in modelling the association between the response variable and the given set of explanatory variables.

3.4.1. Generalized Estimating Equation (GEE)

Generalized estimating equations (GEE) are an extension of GLMs to accommodate correlated data. Generalized Estimating Equations models have known the function of the marginal expectation of the dependent variable as a linear function of one or more explanatory variables. The GEE methodology provides consistent estimators of the regression coefficients and their variances under weak assumptions about the actual correlation among a group's observations. This approach avoids a need for multivariate distributions by assuming only a functional form for marginal distribution at each time point or condition. It relies on independence across groups (regions) to consistently estimate the variance of proposed estimators even when the assumed working correlation structure is incorrect.

For binary data, a GEE approach is used to account for the correlation between responses of interest for subjects from the same cluster[19]. GEE is a non-likelihood method that uses correlation to capture the association within the clusters or subjects in terms of marginal correlations[17, 30].

GEE is a general method for analysing data collected in clusters where

1. Observations(women and daughters) within a cluster are correlated,
2. Women and daughters in separate clusters are independent,
3. A monotone transformation of the expectation is linearly related to the explanatory variables
4. The variance is a function of the expectation[18,20].

For clustered as well as repeated data, Liang and Zeger proposed GEE which requires only the correct specification of the univariate marginal distributions provided one is willing to adopt "working" assumptions about the correlation structure. The "working" assumptions as proposed by Liang and Zeger include independence, exchangeable and auto-regressive AR (1).

Generalized estimating equations have several options to select from in specifying the form of the correlation matrix. This specification will differ based on the nature of the data collected. Although GEE models are generally robust to misspecification of the correlation structure, in cases in which the specified structure does not incorporate all of the information on the correlation of measurements within the cluster, we can expect that inefficient estimators will result. The four common working correlations listed above reviewed below:

An autoregressive: correlation structure is specified to set the within-subject or within-cluster correlations as an exponential function of this lag period, for data that are correlated within cluster over time.

An exchangeable: correlation structure is in which within-subject or within-cluster observations are equally correlated. Where there is no logical ordering for observations within-cluster (such as when data are clustered within the subject or within an organizational unit but not necessarily collected over time), an exchangeable correlation matrix should be used.

Unstructured: working correlation matrix estimates all possible correlations between within-subject or within-cluster responses and includes them in the estimation of the variances. This is free estimation on the within subject correlation from the data.

Independence: is the responses within the subject or cluster are independent of each other; this approach sacrifices one of the two benefits of using GEE in that it does not account for within group correlation but is still useful in model fitting (as a base model)[31].

Note that the independence structure adds about no additional parameters α and hence, when there is no over dispersion, parameter estimates β will not differ from those obtained from logistic regression.

Independence and exchangeable working assumptions can be used in virtually all applications, whether longitudinal, clustered, multivariate, or otherwise correlated. AR(1) and unstructured are less relevant for clustered data, longitudinal studies with unequally spaced measurements and/or sequences with differing lengths. However, even though it seems less advisable to use such structures in cases where they are not supported by the study’s design, it is strictly speaking not a mistake as, once again, working assumptions are allowed to be wrong. Note that the AR(1) parameter is estimated using adjacent pairs of measurements only, in contrast to the exchangeable correlation, for which all pairs within a sequence are employed[29].

3.4.1.1. Generalized Estimating Equations Models

Let $Y_j = \left(Y_{j1} \dots Y_{jnj}\right)^T$ be the response values of observations from jth cluster (from jth region) j = 1, 2,.., m with corresponding vector of means $\mu_j = (\mu_{j1} \dots \mu_{jnj})$ follow a binomial distribution, i.e. $Y_j \sim Bin(n_j, \mu_j)$ that belongs to exponential the family. Correlated data are modelled using the same link function and linear predictor setup (systematic component) as the independence case. The random component is described by the same variance functions as in the independence case, but the covariance structure of the correlated measurements must also be modelled. Let the vector of independent variables for i^{th} an individual is X_{ji}=$[X_{ji1}, \dots, X_{jip}]^T$.

Then to model the relation between the response and covariates, one can use a regression model similar to the generalized linear model (eqn 3.1)

$$logit(\mu_j) = X_j^T \beta \qquad 3.2$$

Where, $logit(\mu_j)$= logit link

X_j= (n_j x P) dimensional vector of known covariates

β = (β_1, β_2... β_p)' (P X 1) dimensional vector of unknown fixed regression parameter to be estimated

$E(Y_j) = \mu_j$ is expected value of responses.

Assume that you have chosen a model that relates a marginal mean to the linear predictor $X_j^T \beta$ through a link function. The generalized estimating equations for estimating β, is an extension of the independence estimating equation to correlated data and is given by:

$$U(\beta, \hat{\alpha}) = \sum_{j=1}^{n} \frac{\partial \mu_j}{\partial \beta}^T V_j^{-1}(y_j - \mu_j) = 0 \qquad 3.3$$

Where V_j is an estimator of the covariance matrix of Y_j and it is specified as the estimator $\hat{\boldsymbol{\alpha}}$ an estimate of the 'nuisance' parameter vector

$$V_j(\alpha) = \varphi A_j^{\frac{1}{2}} R_j(\alpha) A_j^{\frac{1}{2}} \qquad 3.4$$

Where

A_j = is n_j X n_j diagonal matrix with V (μ_{ij}) as i^{th} diagonal element (A_j= diag(V(μ_{ij}))

$R_j(\alpha)$ is n_j x n_j working correlation matrix of within cluster responses that are fully specified by the vector of parameter α.

The i, i' element of $R_j(\alpha)$ is known, hypothesized, or estimated correlation between Y_{ji} and Y_{ji}'. This working correlation matrix may depend on the vector of unknown parameters α, which is the same for all subjects. If $R_j(\alpha)$ is the true correlation matrix of Y_j, then V_j is the true covariance matrix of Y_j.

φ is dispersion parameter and is estimated by $\hat{\varphi} = \frac{1}{N}\sum_{j=1}^{m}\frac{1}{n_j}\sum_{j=1}^{nj} e_{ji}{}^2$

Where N = $\sum_{j=1}^{m} n_j$ is the total number of measurements and

e_{ji} is the Pearson residual given by $e_{ji} = \frac{yji - \mu ji}{\sqrt{V(\mu ji)}}$.

Thus, the score equation used to estimate the marginal regression parameters while accounting for the correlation structure is given by

$$S(\beta) = \sum_{j=1}^{m} \frac{\partial \mu j^T}{\partial \beta}\left[A_j^{\frac{1}{2}} \mathrm{Rj} A_j^{\frac{1}{2}}\right]^{-1}\left(y_{j-}\,\mu_j\right) = 0 \qquad 3.5$$

The model-based estimator of cov($\hat{\beta}$) is given by $\sum_m \hat{\beta}\, I_0{}^{-1}$

Where $I_0 = \sum_{j=1}^{m}\frac{\partial \mu j^T}{\partial \beta} V_j^{-1}\frac{\partial \mu j}{\partial \beta}$

The estimator

$\sum_e = V(\hat{\beta}) = I_0{}^{-1} I_1 I_0{}^{-1}$ is called the empirical, or robust, an estimator of the variance-covariance matrix of $\hat{\beta}$, where $I_1 = \sum_{j=1}^{m}\frac{\partial \mu j^T}{\partial \beta} V_j^{-1} cov\left(y_j^{-1}\right)\frac{\partial \mu j}{\partial \beta}$.

An advantage of the GEE approach is that it yields a consistent estimator of coefficients, even when the working correlation matrix R_j is not specified. However, severe misspecification of working correlation may seriously affect the efficiency of the GEE estimators[29,32].

3.4.1.2. Parameter Estimation

The generalized estimating equations are estimates of quasi-likelihood equations which is quasi-likelihood estimators. A quasi-likelihood estimate of β arises from the maximization of normality-based log likelihood without assuming that the response is normally distributed.

In general, there are no closed-form solutions, so the GEE estimates are obtained by using an iterative algorithm, that is the iterative quasi-scoring procedure. GEE estimates of model parameters are valid even if the covariance is mis-specified (because they depend on the first moment, e.g., mean). However, if the correlation structure is mis-specified, the standard errors are not good, and some adjustments based on the data (empirical adjustment) are needed to get more appropriate standard errors. Points out that a chosen model in practice is never exactly correct, but choosing carefully a working correlation (covariance structure) can help with the efficiency of the estimates[33].

3.4.1.3. Significance Test

We can consider significance tests for individual estimates, such as intercepts, slopes, and their variances, as well as whether the full model accounts for a significant amount of variance in the dependent variable. GEE model parameters are estimated using quasi-likelihood procedures, there is no associated likelihood underlying the model. To determine the significance of the predictor variables we can use Wald statistic from empirical (robust) and model based (naive) estimates. Wald statistics based confidence intervals and hypothesis testing for parameters; recall they rely on asymptotic normality of estimator and their estimated covariance matrix. The likelihood ratio test is not valid in this setting because the estimating equations are not necessarily likelihood equations.

To compare the GEE models; however, one can construct a multi-parameter Wald test to test the null hypothesis that a set of βs equal 0.

Wald test then equals

$$x^2 = \hat{\beta}'C'(CV(\hat{\beta}')C')'C\hat{\beta} \qquad 3.6$$

which is distributed as x^2 with q degrees of freedom under the null hypothesis. The prime symbol ' indicates the transpose of the matrix or vector. Where C is a 1 x p vector selecting a single regression coefficient β [25].

This will help test the hypothesis:

$H_0 = \beta_1 = \beta_2 = \beta_3 = ... = \beta_q$ versus the alternative that

$H_1 = \beta_q \neq \beta_p$, for $q \neq p$

The goodness of Fit Test

Traditional model selection criteria such as AIC or BIC, however, cannot be directly applied for correlation structure selection in GEE because they are likelihood-based and full multivariate likelihoods are not expressed or used in GEE estimation. Instead, the estimation is based (in part) on quasi-likelihood. While defining the full multivariate likelihood is tractable under the assumption of independent responses and in some particular multivariate data settings, but is often not so in the general multivariate, correlated data setting. But, we don't test for the model fit of the GEE, because this is really an estimating procedure; there is no likelihood function. We used some criteria to select a working correlation structure for the GEE methods including quasi-likelihood under the independence model criterion (QIC).

The QIC was constructed by replacing the likelihood in the Kullback-Leibler information with the quasi-likelihood under the working correlation autoregressive, exchangeable and independence assumption. The quasi-likelihood approach is natural in this setting because the quasi-likelihood estimating equations have the same form as the maximum likelihood estimating equations in the GLM-type models when the canonical link is used.

Let M_C be a candidate model and M_T be the true model. The Kullback-Leibler discrepancy between M_C and M_T using the quasi-likelihood under the different correlation model is

$$\Delta(\hat{\beta}, \beta_T, I) = E_{M_T}[-2Q(\beta; I, \Delta) \qquad 3.7$$

Where $Q(\beta; I, \Delta)$ the quasi-likelihood under independence assumption of the dataset Δ and the expectation E_{M_T} is taken under the true model $M_{T.}$

The QIC for GEE is defined as

$$\text{QIC(R)} = -2\text{Q}[\hat{\beta}(R); I, \Delta)] + 2tr[\Sigma_{M(IN)}^{-1} \acute{} \hat{\Sigma}_{S(R)}] \qquad 3.8$$

The estimate $\hat{\beta} = \hat{\beta}(R)$ is the estimated regression parameter vector under the working correlation matrix R of the candidate model M_T.

QIC is free from both the working correlation structure as well as the true correlation structure, so it would not be informative in the selection of the covariance structure[28,34].

To better select “working” correlation structure, Hin and Wang proposed correlation information criterion (CIC) defined by

$$\text{CIC} = \text{tr}(\Sigma_{M(IN)}^{-1} \hat{\Sigma}_{S(R)}) \qquad 3.9$$

In their work, CIC was shown to outperform QIC when the outcomes were binary through simulation studies[34].

3.4.1.4. Model Diagnostic

As with any model, the GEE approach needs diagnostic procedures for checking the model's adequacy, and for detecting outliers and influential observations. However, diagnostic tools such as the likelihood ratio test for generalized linear models are not available for the GEE analysis. Pearson residuals are used to detect extreme observations. The estimated variance-covariance matrix of Pearson residuals is biased since the residuals are correlated to each other.

3.4.2. Multilevel Logistic Regression

In multilevel research, the structure of data in the population is hierarchical, and a sample from that population can be viewed as a multistage sample. Because of different cases like, cost, time and efficiency considerations, stratified multistage samples are the norm for sociological and demographic surveys. For such samples the clustering of the data is, in the part of data analysis and data reporting, a nuisance which should be taken into consideration. Cluster sampling system often introduces multilevel dependency or correlation among the observations that can have implications for model parameter estimates. For multistage clustered samples, the dependence among units or observations often comes from several levels of the hierarchy. The problem of dependencies between individual observations also occurs in survey research, where the sample is not taken randomly but cluster sampling from geographical areas is used instead. In this case, the use of single-level statistical models is not reasonable. Hence, in order to draw appropriate inferences and conclusions from multistage stratified clustered survey data, we may require complicated modelling techniques like multilevel modelling. Multilevel models contain variables measured at different levels of the hierarchy.

3.4.2.1. Multilevel Logistic Regression Model

The multilevel logistic regression analysis considers the variations due to the hierarchy structure in the data. Hence the study examines the effects of group level and individual level variation of observations. Let considering a two-level model for binary outcomes with a single explanatory variable and let y_{ij} be the binary outcome variable, coded "0" or "1", associated with level-one unit i nested within level two unit j. Also let π_{ij} be the probability that the response variable equals 1, $\pi_{ij} = pr(y_{ij} = 1)$. Like the ordinary logistic regression, is modelled using the link function, logit. The two-level logistic regression model can be given as:

$$\ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 x_{ij} + \mu_j \qquad 3.10$$

Where μ_j is the random effect at level 2

Without μ_j, this equation (3.10) can be considered as a standard logistic regression model. Therefore, conditional on μ_j the $y_{ij}{}'s$ can be assumed to be independently distributed. Here μ_j is a random quantity and follows N(0, δ_μ^2). In order to know the variation at each level, the model (3.10) can be written as follows splitting up into two models: one for level 1(individual level) and the other for level 2(regional level).

$$ln\left[\frac{\pi_{ij}}{1-\pi_{ij}}\right] = \begin{cases} \beta_{0j} + \beta_1 x_{ij} & model: level\ 1 \\ \beta_0 + \beta_1 x_{ij} + uj & model: level\ 2 \end{cases} \qquad 3.11$$

That is, β_1 is the difference between the log-odds of the outcome for women in the same cluster in the same division which have observed x values that differ by one unit

Testing Heterogeneity of Proportions

For the proper application of multilevel analysis in general and multilevel logistic regression analysis in particular, the first logical step is to test for heterogeneity of proportions between groups (in our case between Regions). Two commonly used test statistics that are used to check for heterogeneity of proportions are described below.

To test whether there are indeed systematic differences between groups, the well-known chi-square test for contingency tables can be used. The test statistic of the chi-square test for a contingency table is often given in the familiar form:

$$X^2 = \sum\sum\frac{(O_{ij}-E_{ij})^2}{E_{ij}} \qquad 3.12$$

With degrees of freedom equal to numbers of categories minus one i.e. (R - 1) (C - 1)

Where:-

R is row and C is a column

O_{ij} – observed frequency

E_{ij} – expected frequency

It can also be written as:

$$X^2 = \sum_{j=1}^{g} n_j \frac{(\hat{p}_j - \hat{p}_.)^2}{\hat{\hat{p}}_.(1-\hat{\hat{p}}_.)'} \qquad 3.13$$

Where, $\widehat{p_J} = \frac{1}{n_j}\sum_{i=1}^{n_j} Y_{ij}$, the proportion of women who are practicing genital mutilation in region j,

$\widehat{p_J} = \frac{1}{K}\sum_{j=1}^{g}\sum_{i=1}^{n_j} Y_{ij}$, K= $\frac{1}{K}\sum_{i=1}^{n_j} n_j$

Where K is the number of groups, n_j is the number of samples in the j^{th} group. Similarly, the overall average

$\hat{p}_. = \frac{1}{n}\sum_{j=1}^{g}\sum_{i=1}^{n_j} Y_{ij}$ here is the overall proportion of success.

This statistic follows approximately a chi-square distribution with g-1 degrees of freedom. This chi-squared distribution is an approximation valid if the expected number of success ($n_j\hat{p}_j$) and of failures (n_j (1- $\hat{p}_j$)) in each group are at least 1 while 80 percent of them are at least 5[28,35].

Estimation of between and within group variance: the theoretical variance between the group dependent probabilities, i.e., the population value of Var(p_j), can be estimated by:

$$\tau^2 = S^2_{between} - \frac{S^2_{within}}{\tilde{n}} \qquad 3.14$$

Where $\tilde{n}$ is given by:

$$\tilde{n} = \frac{1}{n-1}\left\{n - \frac{\sum_j n_j^2}{n}\right\} = \bar{n} - \frac{S^2 n_j}{g\bar{n}} \quad \text{And}$$

$$n = \sum_{j=1}^{g} n_j$$

For dichotomous outcome variables, the observed between-groups variance is closely related to the chi-squared test statistic (3.13). They are connected by the formula

$$S_{between}^{2} = \frac{\hat{p}_{.}(1-\hat{p}_{.})}{\bar{n}(g-1)} X^{2} \qquad 3.15$$

The within-group variance in the dichotomous case is a function of the group averages, though,

$$S_{within}^{2} = \frac{1}{M-g} \sum_{j=1}^{g} n_j \, \hat{p}_j (1-\hat{p}_j) \qquad 3.16$$

3.4.2.2. The Random Intercept-Only Model

The empty two-level model for a dichotomous outcome variable refers to a population of groups (level-two units, i.e. regions)) and specifies the probability distribution for group dependent probabilities without taking further explanatory variables into account. This model only contains random groups and random variation within groups. It can be expressed with logit link function as follows.

$$logit\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + U_{0j}$$

$$U_{0j} \sim IID(0, \delta_0^2)$$

Where β_0 the population average of the transformed probabilities and U_{0j} is the random deviation from this average of group j.

3.4.2.3. The Random Intercept and Fixed Slope Model

In the random intercept logistic regression model, the intercept is the only random effect meaning that the groups differ with respect to the average value of the response variable. It represents the heterogeneity between groups in the overall response. The logistic random intercept model expresses the log-odds, i.e. the logit of p_{ij}, as a sum of a linear function of the explanatory variables and a random group-dependent deviation U_{0j}. That is,

$$logit(p_{ij}) = \ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \cdots + \beta_k x_{kij}$$

$$logit(p_{ij}) = \beta_{0j} + \textstyle\sum_{h=1}^{k} \beta_h x_{hij} \qquad 3.17$$

Where the intercept term β_{0j} is assumed to vary randomly and is given by the sum of an average intercept β_0 and group dependent deviations U_{0j}. That is:

$$\beta_{0j} = \beta_0 + U_{0j}$$

As a result

$$logit\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \textstyle\sum_{h=1}^{k} \beta_h x_{hij} + U_{0j} \qquad 3.18$$

$$U_{0j} \sim IID(0, \delta_0^2)$$

Where β_0 is the log-odds that y = 1 when x = 0 and u = 0, β_h is the effect on log-odds of the dependent variable in the same group (same value of u), exp (β_h) is an odds ratio, comparing odds for individuals in the same group. U_{0j} is the effect of being in group j on the log-odds that y = 1 also known as a level 2 residual, δ_0^2 is the level 2 (residual) variance, or the between-group variance in the log-odds that y = 1 after accounting for X.

Note that the first part of the left-hand side of (equ 3.18), incorporating the regression coefficients, $\sum_{h=1}^{k} \beta_h x_{hij}$ is the fixed part of the model, because the coefficients are fixed. The remaining part, U_{0j}, is called the random part of the model.

The region intercepts measure the differences between the regions, controlling for other effects in the model. Equation (3.18) is a mixed model because it has both fixed effects and random effects. It is a logistic mixed model, because the link function is logit.

3.4.2.4. The Random Coefficient Model

As yet, we have allowed the probability of female genital mutilation to vary across regions, but we have assumed that the effects of the explanatory variables are the same for each region. We will now modify this assumption by allowing the difference between explanatory variables within a region to vary across regions. To allow for this effect, we will need to introduce a random coefficient for those explanatory variables. So, a random coefficient model represents heterogeneity in the relationship between the response and explanatory variables.

As mentioned above, the response variable in this study, circumcision status was binary. Therefore, the statistical model used in this analysis was the two-level random coefficient multilevel regression model. The model, with p level-1 predictors and q level-2 predictors, can be expressed as follow:

$$\ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \sum_{h=1}^{p} \beta_{hj} x_{hij} + \sum_{h=1}^{q} U_{hj} x_{hj}$$

Where

$\beta_{0j} = \beta_0 + U_{0j}$, i = 1, 2, … n_j, j = 1,2,3, … 11

Now the above equation is written as

$$\text{logit}(\pi_{ij}) = \ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \sum_{h=1}^{p} \beta_{hj} x_{hij} + U_{0j} + \sum_{h=1}^{q} U_{hj} x_{hj} \qquad 3.19$$

$Y_{ij} \sim Binary(\pi_{ij}),\ U_{0j} \sim Normal(0, \delta_0^2)$

The first part of the equation (3.19), $\beta_0 + \sum_{h=1}^{p} \beta_{hj} x_{hij}$ is called the fixed part of the model. The second part $U_{0j} + \sum_{h=1}^{q} U_{hj} x_{hj}$ is called the random part.

Y_{ij} is the value of the response variable for the i^{th} woman in the j^{th} region. X_{hij} is the value of the individual-level explanatory variable for X_p the i^{th} woman in the j^{th} region.

X_{hij} is the value of region-level explanatory variable X_q for the j^{th} region. U_{0j} is the region-level residual of the j^{th} region. It is a random effect that represents the discrepancy between β_0 and the true intercept of the j^{th} region and n_j is the numbers of women respondents in the j^{th} region.

Parameter estimation of fixed effects

The concept of a fixed factor is most commonly used in the setting of a standard analysis of variance (ANOVA) or analysis of covariance (ANCOVA) model. Fixed effects are also known as regression coefficients or fixed-effect parameters, describe the relationships between the dependent variable and predictor variables for an entire population of units of analysis, or for a relatively small number of subpopulations defined by levels of a fixed factor. Fixed effects are assumed to be unknown fixed quantities in a GLMM, and we estimate them based on our analysis of the data collected.

Parameter estimation for the multilevel logistic model is not straightforward like the methods for the simple logistic regression model. In this paper we are focused on hierarchical logistic regression models, which can be fitted using the penalized quasi-likelihood (PQL) and Laplace Approximation of the maximum likelihood method to estimate parameters.

Significance Testing for Fixed Effects

The fixed effects in the multilevel regression are typically tested in a familiar way, by creating a ratio of the intercept or slope estimate to the estimate of the standard error. The usual null hypothesis test is whether the coefficient, either intercept or slope, is significantly different from zero (i.e., is the population value zero or not). This kind of ratio, usually distributed as a z or t, is used in many statistical tests.

$$t = \frac{\hat{\beta}_h}{S.E(\hat{\beta}_h)}$$

Where $\hat{\beta}_h$ is either the intercept or slope coefficient and $S.E(\hat{\beta}_h)$ is the standard error estimate.

The R software will be required to estimate fixed effects tests involve the same ratio of the estimate to the standard error estimate, but the significance is determined by the normal curve, so it is considered a z-test. The z-test is often referred to as a Wald test.

Random effects

Random effects are random values associated with the levels of a random factor (or factors) in a GLMM. These values, which are specific to a given level of a random factor, usually represent random deviations from the relationships described by fixed effects. For example, random effects associated with the levels of a random factor can enter a GLMM as random intercepts (representing random deviations for a given subject or cluster from the overall fixed intercept), or as random coefficients (representing random deviations for a given subject or cluster from the overall fixed effects) in the model. In contrast to fixed effects, random effects are represented as random variables in a GLMM.

Significance Testing for Random Effects

Random effects tests examine hypotheses about whether the variance of intercept or slopes (or their covariance) is significantly different from zero. The tests of variances and covariance's are made using a likelihood ratio test and deviance based chi-square test.

The goodness of Fit Test

Another approach to significance tests involves a comparison of two-nested models. Nested model tests involve the comparison of one model to another model that specifies only a subset of the parameters included in the first model (provided the same set of cases are used in both models).

The likelihood ratio test compares the deviance (-2 log likelihood) of two models by subtracting the smaller deviance (model with more parameters) from the larger deviance (model with larger deviance). The difference is a chi-square test with the number of degrees of freedom equal to the number of different parameters in the two models. Any number of parameters can be compared in the two models. In the case where the empty model is compared to a full model, the likelihood ratio test provides information about whether the predictors in the model together account for a significant amount of variance in the dependent variable.

AIC and BIC

The AIC (Akaike Information Criterion) and the BIC (Bayesian Information Criterion) are two popular measures for comparing maximum likelihood models. AIC and BIC are defined as:

$$AIC= -2*\ln(\text{Likelihood}) + 2*K$$

$$BIC= -2*\ln(\text{Likelihood}) + \ln(N)*K$$

Where k is the model degrees of freedom calculated as the rank of variance–covariance matrix of the parameters and N is the number of observations used in estimation or, more precisely, the number of independent terms in the likelihood. AIC and BIC can be viewed as measures that combine fit and complexity. The fit is measured negatively by -2*ln (Likelihood). The larger the value, the worse the fit is. Complexity is measured positively, either by 2*k(AIC) or ln (N)*k(BIC). The larger the value also, the worse the fit model. Given two models fit on the same data, the model with the smaller value of the information criterion is considered to be better[36].

3.4.2.5. Model Diagnostic

It is of interest to obtain the residual values from the estimated multilevel model. Plots are a good way to examine the residuals. But in multilevel logistic regression, many different residual plots can be used. For this study, the fitted model was checked for the possible presence of outliers and influential values in a similar fashion with the standard logistic model. But additionally, the presence of outliers and influential observation were examined for level two. Leverage and influence value greater than one is considered as an influential observation for both level one and level two.

CHAPTER FOUR

4. RESULT AND DISCUSSION

The results of the study are discussed in this chapter. The response variable, female genital mutilation is categorical. The genital mutilation status is 1 for circumcised women and daughters, 0 for not. In this study, GEE and the two-level logistic regression model are used to see the relationship between the proposed independent variables and the response variables. We start our data analysis by descriptive statistics for the categorical variables considered in the study.

4.1. Descriptive Statistics

The data used for this study was extracted from EDHS in 20116. The data contains 15683 women respondents. All the variables in this study are categorical. Summary of the response variables is presented in Table 1A and Table 1B in Appendix B.

About 72.6% of women in this study responded to the question of circumcision of themselves (25.8% and 46.8% of response were, “NO” and “YES” respectively). About 27.4% of response for this variable is missing.

When women asked about the circumcision status of their daughters nearly 36% did not respond. This number is high when compared to circumcision of themselves, even when the woman does not have daughters (Table 4B in Appendix B).

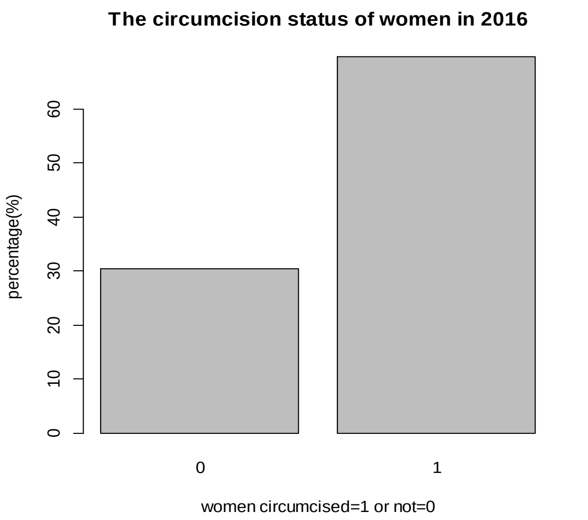


Figure 4. 1a: The circumcision status of Women

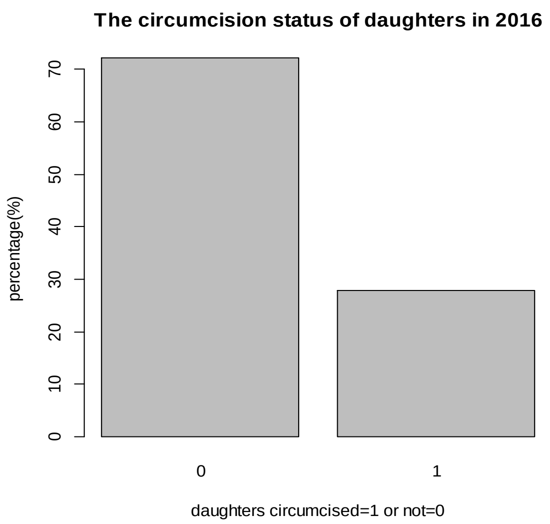


Figure 4.1b: The circumcision status of daughters

The figure 4.1a show that 70.4% of women in the age group 15-49 in Ethiopia are circumcised in 2016. Reversely the mother who circumcises their daughters is only 22.7%.

Prevalence of Female Genital Mutilation

Table 4.1. 1: The prevalence of FGM among women (15-49 ages) and daughters (<15 ages) circumcision by mothers cultural, socio-economic and demographic factors:

Explanatory Variables	Weighted Number of women circumcised	Weighted Number of Women	Percentage of women Circumcised (%)	Weighted Number of women circumcise their daughters	Weighted Number of daughters	Prevalence of daughter Circumcised (%)
Region						
Tigray	131	462	28.3	31	219	14.1
Afar	60	66	91.2	27	36	75
Amhara	1127	1702	66.2	358	859	41.7
Oromia	2178	2747	79.3	199	1589	12.5
Somali	225	228	98.6	58	142	40.8
Benishangul	47	68	69.1	9	39	23
SNNPR	1024	1432	71.5	175	788	22.2
Gambela	7	16	46.3	0.5	6	8.3
Harari	15	18	82.3	1	10	10
Addis Ababa	251	461	54.5	2	101	2
Dire Dawa	35	45	77.4	2	17	11.7
Total	5101	7249	70.4	867	3806	22.7
Residence						
Urban	923	1665	55.5	50	534	9.3
Rural	4177	5583	74.8	817	3272	24.9
Age in five years group						
15-19	786	1523	51.6	4	81	4.9
20-24	756	1194	63.3	48	412	11.6
25-29	996	1356	73.4	140	883	15.8
30-34	923	1124	82.1	199	910	21.8

35-39	758	939	80.7	225	740	30.4
40-44	471	610	77.2	121	449	26.7
45-49	409	502	81.5	127	329	38.6
Religion						
Orthodox	1856	3158	58.7	407	1456	27.9
Catholic	38	52	73	9	27	33.3
Protestant	1225	1674	73.2	181	897	20.1
Muslim	1941	2288	84.8	269	1371	19.6
Traditional	34	55	61.8	0	37	0
Other	5	22	22.7	0	16	0
Wealth Index						
Poorest	851	1127	75.5	188	738	25.4
Poorer	939	1284	73.1	176	779	22.6
Middle	1086	1408	77.1	191	798	23.4
Richer	1062	1445	73.5	226	809	27.9
Richest	1160	1984	58.4	84	682	12.3
Women Education Status						
No Education	2759	3406	81	688	2539	27
Primary	1661	2505	66.3	169	970	17.4
Secondary	452	889	50.8	8	192	4.1
Higher	227	449	50.5	0	104	0
Mother circumcision status						
Yes				834	3041	27.4
No				33	765	4.4
Media exposure						
Yes				232	1333	17.4

No	636	2473	25.7
Father Education level			
No Education	530	1759	30.1
Primary	209	1208	17.3
Secondary	29	249	11.7
Higher	13	198	6.7
Don't know	5	16	31.6

The prevalence of FGM/C among women was 70.4% in 2016, which is obtained after data were weighted. Women in rural areas are more likely to be circumcised than women in urban areas (74.8% and 55.5%, respectively). The prevalence rate of FGM among women also vary across the region. The highest prevalence was recorded in Somali region by 98.6% and followed by the Afar region (91.2%). Tigray has the lowest prevalence (28.3%), followed by Gambela (43.7%). The prevalence of FGM/C increases with age, 51% among women age 15-19, 63.3% among women age 20-24, 73.4 percent among women age 25-29, 82.1% among women age 30-34 and almost the same prevalence among the oldest women (30-49 ages). The prevalence of circumcised women is lowest among Orthodox women (58.7%) and highest among Muslim women (84.8%). And the FGC is less prevalent among women with higher education and those in the highest wealth quintile.

The prevalence of FGM/C among daughters was 22.7%, and 24.9% of rural daughters are circumcised, as compared with 9.3% of daughters in urban areas. The prevalence of FGM among daughters also showed regional variability. The highest prevalence was recorded in Afar region by 75% and less prevalence recorded in Addis Ababa by two percent. The result shows the proportion of FGM/C among daughter's increases, as their mother ages increases (4.9% among daughters of their mother age was 15-19, 11.6% among daughters of their mother age was 20-24, 15.8% among daughters of their mother age was 25-29, 21.8% among daughters of their mother age was 30-34 and 30.4%, 26.7% , 38.6% among daughters of their mother age was 35-39, 40-44, 45-49 respectively). The percentage of daughter's genital mutilation in urban and rural is 9.3% and 24.9% respectively. The prevalence of female genital mutilation among daughters of mothers who has no education, primary, secondary, higher is 27%, 17.4%, 4.1%, 0% respectively. This shows us female genital mutilation among daughters is decrease, as their mother education level is increased. The prevalence of female genital mutilation among daughters of their father have no education, primary, secondary and higher is 30.1%, 17.3%, 11.7%, and 6.7% respectively. This shows us female genital mutilation among daughters is decreased, as their father education level is increased. 27.4% of daughters from their mother had gone FGM are circumcised, whereas, 4.4% of daughters from their mother not circumcised.

4.2. Generalized Estimating Equations

In this section we fit a model, select the appropriate one and estimate parameters average response over the population (population-averaged effects). In order to maintain the original respondents, the missing data points were imputed randomly using random hot-deck multiple imputations. The R software was used to analyse the data.

4.2.1. Model Building and Selection

The Generalized Estimating Equations models were used to identify the basic factors related to female genital mutilation or cutting at the national level. The models fitted using different correlation structure (autoregressive, exchangeable and independence) were compared by Quasi Information Criteria (QIC) and using CIC. We had also tried to consider unstructured correlation structure, but the Hessian matrix did not converge (or cannot allocate vector large size). In order to select the important factors related to the female genital mutilation of women and daughters, the backward selection procedure was used.

Selection of Correlation Structure

Turning first to the main effects model or model selection, Table 4.2.1 shows the results of the QIC and CIC of the GEE model with autoregressive, exchangeable and independence correlation structure. QIC and CIC are used in selecting a working correlation structure under GEE settings.

Table 4.2. 1: A model selection from the correlation structure of auto-regressive, independence and exchangeable. (For women FGM)

	Autoregressive (ftwar)	**Independence (ftwind)**	**Exchangeable (ftwex)**
QIC	18521.3	18521.6	18523.1
CIC	30.9	31.1	31.2

For the GEE model using autoregressive correlation structure, the QIC is 18521.3 and CIC is 30.9. And the models with independence and exchangeable correlation structure their QIC were 18521.6 and 18523.1 respectively. But, as we discussed in chapter three, autoregressive correlation structure is less relevant in clustered data. Then, the next smallest QIC and CIC is the model fitted with independence correlation structure. Thus, we can conclude that model ftwind is the best fit, and our subsequent discussion focuses on this model. After comparison of the model by their QIC of correlation structure, the next step is a selection of variance of parameter estimates (robust and model based standard error).

As a commonly practiced, comparison of sandwich (empirical) and model based standard errors for the parameter estimates were found in this study. Sandwich and jack-knife variance estimators are a common tool used for variance estimation of parameter estimates of the selected correlation structure. The most efficient statistical inference in GEE would be to use the model-based variance estimator under the correctly specified working correlation structure[37]. This is consistent with our model of GEE for women of reproductive age. Thus, jack-knife (model based) standard error is less compared with robust standard error, which is the standard error used for parameter estimates. That showed on table 4.2.2.

Results of Generalized Estimating Equations

For this particular data, we used the independence correlation structure for the final analysis. Our proposed final model from the model for women circumcised response was given as below excluding wealth index because of the large p-value:

$$\text{logit}(\hat{\mu}_{ji}) = \hat{\beta}_0 + \hat{\beta}_1 regi._{2i} + \hat{\beta}_2 regi._{3i} + \hat{\beta}_3 regi._{4i} + \cdots + \hat{\beta}_{10} regi._{11i} + \hat{\beta}_{11} rural_i + \hat{\beta}_{12} agec_{2i} + \hat{\beta}_{13} agec_{3i} + \hat{\beta}_{14} agec_{4i} + \hat{\beta}_{15} agec_{5i} + \hat{\beta}_{16} agec_{6i} + \hat{\beta}_{17} agec_{7i} +$$

$\hat{\beta}_{18}edsta._{2i}+\hat{\beta}_{19}edsta._{3i}+\hat{\beta}_{20}edsta._{4i}+\hat{\beta}_{21}relg._{2i}+\hat{\beta}_{22}relg._{3i}+\hat{\beta}_{23}relg._{4i}+$
$\hat{\beta}_{24}relg._{5i}+\hat{\beta}_{25}relg._{6i}.$

where reg.2=Afar, reg.3=Amhara, reg.4=Oromoia, reg.5=Somali, reg.6=Ben-Gumuz, reg.7= SNNP,reg.8=SNNP, reg.9=Harari, reg.10=Addis Ababa,reg.11=Dire Dawa, agec2=20-24, agec3=25-29, agec4=30-34, agec5=35-39, agec6=40-44, agec7=45-49, edsta.2=primary, edsta.3=secondary, edsta.4=higher, relg.2 = Catholoic, relg.3 = protestant, relg.4 = Muslim, relg.5 = Traditional, relg.6 = other(Jehovah, etc) and $\mu_{ji}$is the population average the circumcised women.

Table 4 in Appendix B shows that all variables except wealth status were found statistically significant for all models with considered three working correlations. The results suggest significant regional variations in female genital mutilation or cutting in Ethiopia.

Model of GEE with independence correlation structure model in Table 4.2.2 shows that, the risk of FGM is 2.806 times higher among women in Somali as compared to women in Tigray region. The circumcision is 2.778 times higher among women in Afar when compared to among women in Tigray region. This model also shows that the risk of circumcision is 2.234 times and 2.131 times higher in Harari and Dire Dawa region respectively as compared to circumcision in the Tigray region. Gambela region is found to have the lower risk of genital mutilation compared with other regions and was 35.1 percent higher than Tigray region.

The place of residence has a significant effect on female genital cutting/mutilation. The FGM risk is 24.3 percent higher in rural women when compared to urban women.

The result of the selected model provided in table 4.2.2 shows that circumcision is 1.7 times higher among women in 45-49 age groups as compared to among women in 15-19 age groups. And also circumcision is 1.675 times higher among women 35-39 age group as compared to among women in 15-19 age groups. The result also shows that education level has significant effect on female circumcision. The genital mutilation shows decline as the education level of women (respondent) goes increases. The women circumcision is 31.4 percent lower among women in higher education when we compare to those women have no education.

Religion also the factor significantly related to genital mutilation as the result indicates. Muslim women had 1.6 times higher odds of having experienced FGM as compared to

Orthodox women. The circumcision of women of catholic religion follower also shows 1.6 times higher as compared to among women of the orthodox follower. Protestant, traditional and other religions are insignificant at 5% level of significance.

Table 4.2. 2 Parameter estimates, empirical standard errors (sandwich se) and model-based standard errors (jack-knife se) for GEE of women circumcised.

Variables/ Predictors	GEE, with independence					GEE, with exchangeable				
	Est (expβ)	San. se	fij	95% CI	Pr(>|W|)	Est (expβ)	San. se	Fij	95% CI	Pr(>|W|)
(Intercept)	-0.378(0.68)	0.132	0.131	(-0.634, -0.121	0.004	-0.400(0.67)	0.130	0.129	(-0.655, -0.145	0.002
Region(Tigray)	1					1				
Afar	1.022(2.778)	0.143	0.144	(0.739, 1.304)	0.000	1.022(2.778)	0.145	0.147	(0.738, 1.31)	0.000
Amhara	0.568(1.764)	0.104	0.104	(0.364, 0.772)	0.000	0.560(1.750)	0.106	0.106	(0.352, 0.768)	0.000
Oromia	0.628(1.873)	0.105	0.105	(0.422, 0.833)	0.000	0.615(1.849)	0.106	0.105	(0.407, 0.823)	0.000
Somali	1.032(2.806)	0.111	0.111	(0.814, 1.250)	0.000	1.029(2.798)	0.112	0.111	(0.809, 1.25)	0.000
Ben-Gumuz	0.406(1.500)	0.123	0.124	(0.165, 0.647)	0.001	0.431(1.538)	0.122	0.122	(0.192, 0.67)	0.000
SNNP	0.592(1.807)	0.124	0.124	(0.349, 0.835)	0.000	0.576(1.778)	0.123	0.123	(0.335, 0.817)	0.000
Gambela	0.301(1.351)	0.103	0.103	(0.099, 0.503)	0.004	0.292(1.339)	0.105	0.106	(0.086, 0.498)	0.006
Harari	0.804(2.234)	0.117	0.117	(0.575, 1.030)	0.000	0.804(2.234)	0.118	0.118	(0.573, 1.04)	0.000
A/Ababa	0.595(1.813)	0.093	0.094	(0.413, 0.777)	0.000	0.604(1.829)	0.094	0.094	(0.419, 0.788)	0.000
D/Dawa	0.757(2.131)	0.105	0.105	(0.551, 0.963)	0.000	0.766(2.151)	0.106	0.106	(0.558, 0.973)	0.000
Residence(Urban)	1					1				
Rural	0.218(1.243)	0.092	0.092	(0.037, 0.398)	0.017	0.251(1.285)	0.089	0.089	(0.076, 0.425)	0.005
Age cate.(15-19)	1					1				
20-24	0.303(1.353)	0.057	0.055	(0.191, 0.415)	0.000	0.297(1.345)	0.057	0.055	(0.185, 0.408)	0.000
25-29	0.340(1.404)	0.060	0.059	(0.222, 0.458)	0.000	0.333(1.395)	0.061	0.059	(0.213, 0.452)	0.000

30-34	0.517(1.676)	0.070	0.068	(0.380, 0.654)	0.000	0.515(1.673)	0.070	0.068	(0.377, 0.652)	0.000
35-39	0.516(1.675)	0.073	0.071	(0.373, 0.659)	0.000	0.507(1.660)	0.072	0.070	(0.365, 0.648)	0.000
40-44	0.412(1.509)	0.079	0.077	(0.257, 0.567)	0.000	0.431(1.538)	0.079	0.077	(0.276, 0.585)	0.000
45-49	0.530(1.698)	0.088	0.086	(0.358, 0.702)	0.000	0.527(1.693)	0.087	0.085	(0.356, 0.697)	0.000
Educ. Status (No Edu)	1					1				
Primary	-0.051(0.95)	0.053	0.052	(-0.155, 0.052)	0.329	-0.043(0.95)	0.051	0.050	(-0.142, 0.056)	0.393
Secondary	-0.286(0.75)	0.069	0.067	(-0.421,-0.151)	0.000	-0.269(0.76)	0.067	0.066	(-0.400, -0.137)	0.000
Higher	-0.379(0.68)	0.085	0.084	(-0.546, -0.212	0.000	-0.362(0.69)	0.084	0.083	(-0.526, -0.197	0.000
Religion(Orthodox)	1					1				
Catholoc	0.442(1.555)	0.221	0.226	(0.009, 0.875)	0.050	0.431(1.538)	0.228	0.233	(-0.015, 0.877)	0.065
Protestant	0.039(1.039)	0.078	0.078	(-0.114, 0.192)	0.613	0.032(1.032)	0.073	0.072	(-0.111, 0.175)	0.660
Muslim	0.474(1.606)	0.066	0.065	(0.345, 0.603)	0.000	0.457(1.579)	0.065	0.064	(0.329, 0.584)	0.000
Traditional	0.068(1.070)	0.248	0.266	(-0.418, 0.554)	0.799	0.144(1.154)	0.272	0.288	(-0.389, 0.677)	0.618
Other	-0.211(0.80)	0.265	0.271	(-0.73, 0.308)	0.436	-0.134(0.87)	0.265	0.267	(-0.653, 0.385)	0.617

NB: san.se = sandwich standard error, fij = fully iterated jack-knife standard error.

The result of Table1 2 in Appendix B shows the model selection of the model fitted for the daughter's genital mutilation. The QIC and CIC results for the model with autoregressive and independence correlation structure are equal, which is small compared to model with exchangeable correlation. But as discussed in chapter three above, the autoregressive working correlation is less relevant in clustered data. So that model with independence correlation structure is the selected one in this study. Depend on the selected correlation structure; most of the parameters with the sandwich standard errors have a less standard error than jack-knife. Therefore we used sandwich standard error for the parameter estimates of the GEE model for daughter's circumcision.

Similar to performed for women circumcision model fit above, we used back ward selection here to identify the potential factors related to female genital mutilation among daughters. Regarding this, the only region, the age of mother and the circumcision status of mother are significant. The type of residence, religion, level of the mother education, level of father education, wealth index and media exposure were does not significantly related to the circumcision of daughters in Ethiopia.

The result of Table 4.2.3 showed that, the circumcised daughters from mother of age groups 20-24 and 25-29 were 14.0% and 18.9% times lower than as compared to circumcised daughters from mother of age groups 15-19, respectively. The estimated odds of genital mutilation for daughters among their mother 44-49 age groups is 1.182 times the estimated odds for daughters among their mother 15-19 age group. The circumcised daughters from Afar, Somali, Amhara and Benishangul-Gumuz regions were 2.5, 1.6, 1.5 and 1.3 times lower than as compared to circumcised daughters from Tigray region, respectively. And the estimated odds of circumcision for daughters among their mother were circumcised are 1.2 times the estimated odds of circumcision for daughters among their mothers were circumcised.

Table 4.2. 3: Parameter estimates, empirical standard errors (sandwich se) and model-based standard errors (jack-knife se) for GEE of daughters circumcision.

Variables/ Predictors	GEE, with independence					GEE, with exchangeable				
	Est (expβ)	San. se	fij	95% CI	Pr(>|W|)	Est (expβ)	San. se	Fij	95% CI	Pr(>|W|)
(Intercept)	-1.311(0.269	0.076	0.077	(-1.459, -1.162)	0.000	-1.308(0.270	0.075	0.076	(-1.455, -1.161)	0.000
Regio(Tigray)	1					1				
Afar	0.908(2.479)	0.099	0.100	(0.713, 1.102)	0.000	0.911(2.486)	0.099	0.100	(0.716, 1.105)	0.000
Amhara	0.384(1.468)	0.096	0.097	(0.195, 0.572)	0.000	0.384(1.468)	0.095	0.096	(0.197, 0.570)	0.000
Oromia	-0.021(0.98)	0.092	0.093	(-0.201, 0.159)	0.816	-0.021(0.98)	0.092	0.093	(-0.201, 0.159)	0.823
Somali	0.467(1.595)	0.091	0.092	(0.288, 0.645)	0.000	0.465(1.592)	0.090	0.091	(0.288, 0.641)	0.000
Ben-Gumuz	0.255(1.291)	0.097	0.098	(0.064, 0.445)	0.008	0.256(1.291)	0.096	0.097	(0.067, 0.444)	0.008
SNNP	0.231(1.26)	0.094	0.095	(0.046, 0.415)	0.014	0.232(1.261)	0.093	0.094	(0.049, 0.414)	0.013
Gambela	0.053(1.055)	0.109	0.111	(-0.160, 0.266)	0.624	0.053(1.054)	0.108	0.110	(-0.158, 0.264)	0.627
Harari	0.044(1.044)	0.116	0.118	(-0.183, 0.271)	0.704	0.045(1.046)	0.116	0.118	(-0.182, 0.272)	0.694
A/Ababa	0.104(1.110)	0.092	0.093	(-0.076, 0.284)	0.260	0.101(1.106)	0.092	0.092	(-0.079, 0.281)	0.270

D/Dawa	0.118(1.125)	0.093	0.094	(-0.064, 0.300)	0.203	0.119(1.126)	0.092	0.093	(-0.061, 0.299)	0.198
Age cate.(15-19)	1					1				
20-24	-0.140(0.86)	0.057	0.056	(-0.251, -0.028)	0.014	-0.138(0.871	0.057	0.056	(-0.249, -0.026)	0.015
25-29	-0.208(0.81)	0.059	0.059	(-0.323, -0.092)	0.000	-0.208(0.81)	0.059	0.059	(-0.323, -0.092)	0.000
30-34	-0.042(0.959	0.063	0.062	(-0.165, 0.081)	0.509	-0.041(0.959	0.063	0.062	(-0.164, 0.082)	0.514
35-39	0.043(1.044	0.063	0.062	(-0.080, 0.166)	0.489	0.044(1.044)	0.063	0.062	(-0.079, 0.167)	0.479
40-44	0.102(1.107)	0.078	0.077	(-0.050, 0.254)	0.190	0.106(1.111)	0.078	0.077	(-0.046, 0.258)	0.172
45-49	0.167(1.182)	0.078	0.077	(0.014, 0.319)	0.031	0.164(1.178)	0.078	0.077	(0.011, 0.316)	0.034
Mother circumcision status (not circ.)										
circumcised	0.186(1.204)	0.044	0.044	(0.099, 0.272)	0.000	0.181(1.198)	0.044	0.044	(0.094, 0.267)	0.000

4.2.2. GEE Model Diagnostic

After model fitting, the next in model building is to perform an analysis of residuals and diagnostics to study the influence of observations and taking appropriate remedial measures. A failure to detect outliers and hence influential observations can have severe distortion on the validity of the inferences drawn from a model. The residuals like Pearson and standardized residuals are also used in order to check model diagnostic. Figure 4.2 and plots under Appendix C did not show any systematic pattern. This indicates that the model fits the data well.

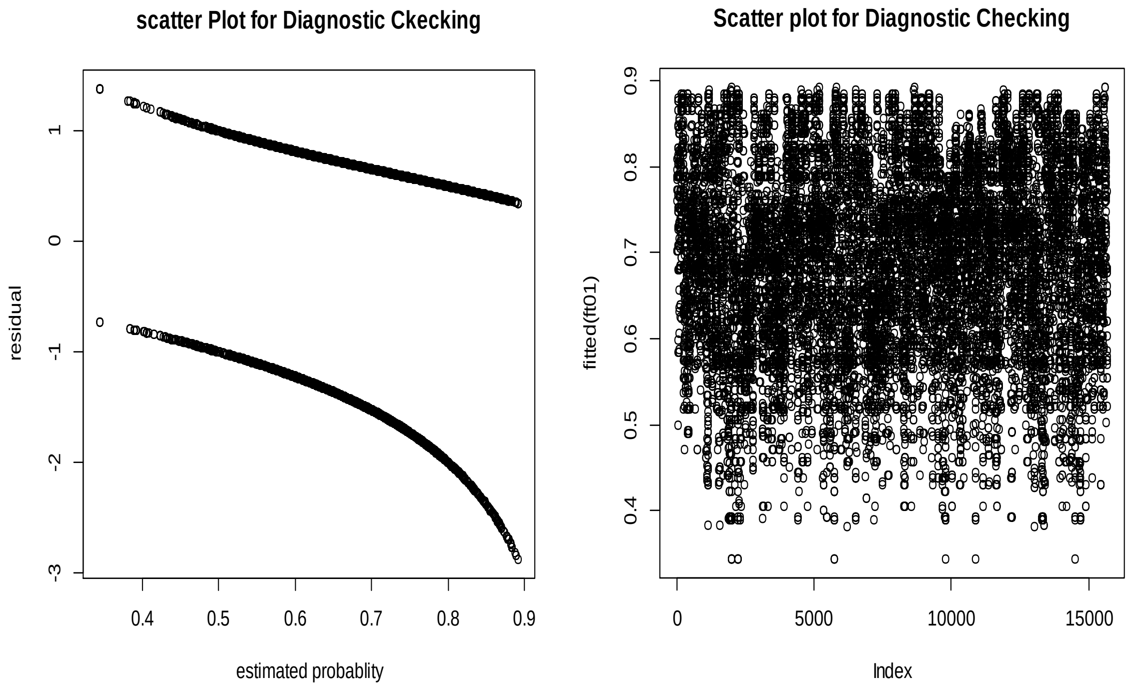


Figure 4. 2: The Plot of the Model Diagnostic Checking

4.3. Multilevel Analysis

We used the penalized quasi-likelihood (PQL) and Laplace Approximation of maximum likelihood method to estimate parameters. The estimated parameters using the two methods are almost the same. The penalized quasi-likelihood estimation provide within group (with in region) variance in addition to between region, whereas the adaptive gauss-Hermite quadrature estimation (Laplace Approximation) method provide only between region variance in random part. But, the maximum likelihood method estimation using penalized quasi-likelihood doesn't provide AIC and analyses of variance (ANOVA) in order to test the significance of the variable, whereas the second method provides those clearly using R software.

4.3.1. The Intercept-only two-level model

We first consider a random intercept only model in order to examine the variation due to the regional effects. We began by fitting a null two-level model that is a model with only an intercept and regional effects.

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + u_{0j}$$

The Intercept β_0 is shared by all regions while the random effect u_{0j} is specific to region j. The random effect is assumed to follow a normal distribution with variance $\delta^2_{u_0}$[36].

From the model estimates presented in Table 4.3.1, we can say that the log-odds of circumcised women in an average region (one with u_{0j}) is estimated as $\widehat{\beta_0}$ = 0.8721. The intercept for region j is 0.8721+u_{0j}, where the variance of $u_{0j}$is estimated as $\delta^2_{u_0}$ = 0.1922.

The likelihood ratio statistic for testing the null hypothesis, that $\delta^2_{u_0}$= 0, can be calculated by comparing the two-level model, with the corresponding single-level model or the standard logit model. Depend on the logLik = -9418.0 of two-level model from table 4.3.1 and logLik = -9686.245 of single-level model from result of table 3 in appendix B, the difference is logLik(single-level) – logLik(two-level) = -268.2682. The test statistic is (X^2 =536.54 (-2*(-268.2682))) with 1 degree of freedom and reject the null hypothesis at 0.05 level of significance, so there is strong evidence that the between region variance is non-zero. This shows that the circumcision status of women is varying among regions.

Table 4.3. 1. The null (random intercept only) model of the two levels on female genital mutilation of women

Fixed Effect	**estimate**	**se**	**Exp(β)**	**95%CI**	**p-Value**
$\boldsymbol{\beta_0}$(constant)	0.8721	0.1335	2.3919	(0.61,1.13)	6.47e-11
Random effect					
$\boldsymbol{\delta^2_{u_0}}$ (between regions)	0.1922	0.4384			
AIC = 18840.0	BIC = 18855.3	Loglik = -9418.0	Deviance = 18836		

Now we examine estimates of the region effects or residuals, $\hat{u}_{0j}$ obtained from the null model. To calculate the residuals and produce a caterpillar plot with the region effects shown in rank order together with 95% confidence intervals.

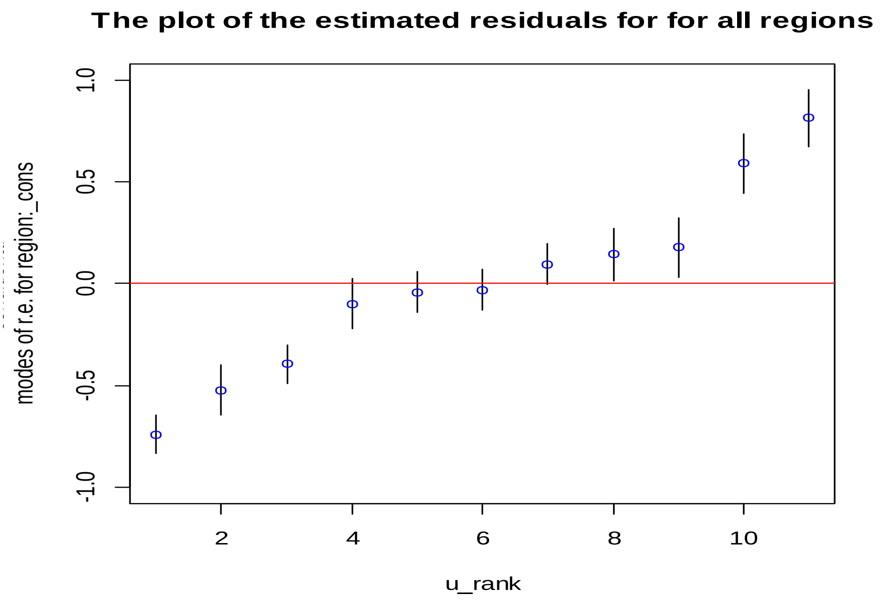


Figure 4. 3: The Plot of the estimated residuals (random effects) for all regions (case of women)

The plot 4.2 shows the estimated residuals for all 11 regions in the sample. For a significant number of regions, the 95% confidence interval does not overlap the horizontal line at zero, that indicates the circumcision status of women in these regions is significantly above average (above the zero line) or below average (below the zero line).

4.3.2. The Random Intercept and Fixed Slope model for women genital mutilation

The next step in model fitting with this data is to add explanatory (predictor) variables in order to identify their relation to the response variable using multilevel logistic regression model. Various comparative statistics of the random intercept-only model (empty model) versus the random intercept and fixed slope model are shown in Table 4.3.2. The random intercept model with fixed explanatory variables is a better fit as compared to the empty model, because of the AIC of random intercept model (18546.9) is less as compared to that AIC of random intercept only model (18840.0).

Table 4.3. 2: Comparisons of the AIC between the intercept-only random model, and random intercept and fixed slope models regarding female genital mutilation of women cases.

Model	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
fit00(Empty two-level model)	2	18840	18855	-9418	18836			
fit0(Saturated two-level model)	17	18550	18680	-9258	18516	142	1	<2e-16 ***

*** Significant (P-value <0.05)

After the selection of model we went to identify the significant variables using back-ward selection. Concerning that wealth index variable is not significant. Omitting wealth index variable we fit again to produce Table 4.3.3.

As can be seen from the table, the likelihood ratio statistic for testing the null hypothesis, that $\delta_{u_0}^2 = 0$, can be calculated by comparing the two-level of the saturated model with a single level of the saturated model. That -2[logLik(two-level saturated model) - logLik(single-level saturated model)] = -2[-9329.045 - (-9257.9)] is 142.29(-2*(-71.145)) (df=16), which is significant at 5% significance level. Which implies the circumcision status of women is varying among regions. In addition, place of residence; educational levels, age and religion were having a significant relation of female genital mutilation among the regions (p-values< 0.05).

Controlling for the effects of other variables and allowing the intercept parameter to vary across regions, the likelihood of women circumcised in rural is about 1.18 times higher as compared to urban women. The likelihood of circumcised women from age groups 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49, were 1.209, 1.38, 1.58, 1.55, 1.44 and 1.66 times higher as compared to women from 15-19 age groups, respectively. Muslim women had 1.64 times higher odds of having experienced FGM as compared to Orthodox women. The women circumcision were 18.4% and 34.0% lowers among women in secondary and higher education respectively, when we compare to those women have no education.

Table 4.3. 3: Results of the logistic multilevel model with a random intercept for the women characteristics on female genital mutilation.

Fixed effect		**Estimate**	**Se**	**Exp(β)**	**95% CI**	**p-value**
	β_0(constant)	0.330963*	0.1124	1.3923	(0.11,0.55)	0.003256
Place of residence	Urban(reference			1		
	Rural	0.170071*	0.0521	1.1853	(0.06,0.27)	0.001116
Age	15-19(ref.)			1		
	20-24	0.189964*	0.0554	1.2092	(0.08,0.29)	0.000615
	25-29	0.329102*	0.0578	1.3897	(0.21,0.44)	1.32e-08
	30-34	0.462004*	0.0646	1.5872	(0.33,0.58)	8.95e-13

	35-39	0.440432*	0.0676	1.5533	(0.30,0.57)	7.39e-11
	40-44	0.367385*	0.0763	1.4439	(0.21,0.51)	1.49e-06
	45-49	0.510844*	0.0871	1.6666	(0.34,0.68)	4.56e-09
Education	No education(ref.)			1		
	Primary	0.006211	0.0484	1.0062	(-0.08,0.10)	0.897962
	Secondary	-0.20362*	0.0626	0.8157	(-0.32,-0.08)	0.001162
	higher	-0.40766*	0.0769	0.6652	(-0.55,-0.25)	1.18e-07
Religion	Orthodox(ref.)			1		
	Catholic	0.360856	0.2383	1.4345	(-0.10,0.82)	0.130007
	Protestant	0.071859	0.0617	1.0745	(-0.04,0.19)	0.244725
	Muslim	0.498196*	0.0544	1.6457	(0.39,0.60)	< 2e-16
	Traditional	-0.242412	0.2305	0.7847	(-0.69,0.20)	0.293036
	Other	0.193534	0.2627	1.2135	(-0.32,0.70)	0.461453
Random Effect						
Level(regional effects)		$\delta^2_{u_0}$	0.0803			
AIC = 18549.8 BIC = 18680.0 logLik = -9257.9 deviance = 18515.8						
Ho: $\boldsymbol{\delta^2_{u_0}}$ = 0, $\boldsymbol{X^2_{calc}}$= 142.29(df=16) logLik(single level) = -9329.045						

Note: ref. = Reference categories. * Significant (P-value <0.05)

4.3.3. Multilevel logistic regression analysis of daughter's circumcision

In this study we also examine the circumcision status of daughters that their ages were less than 15, depending on the responses of women. As the above analysis of women circumcision case, we also apply two level analyses. The random intercept-only model, random intercept and fixed slope model and random coefficient model are compared to fit the best model.

4.3.4. Random Intercept-only Model for Daughters Genital Mutilation

The likelihood ratio statistic for testing the null hypothesis, that $\delta^2_{u_0}$= 0, can be calculated by -2*(logLik(fitted model of single level)-logLik(fitted model of two level))= 166.7113 (df=1) which is large and we reject the null hypothesis at 0.05 level of significance. This indicates that, the variance between regions is non zero regarding daughters case of female genital mutilation.

From the model estimates presented in Table 4.3.4, we can say that the log-odds of circumcised daughters in an average region (one with u_{0j}) is estimated as $\widehat{\beta_0}$ = -0.9338. The intercept for region j is -0.93382 + u_{0j}, where the variance of u_{0j} is estimated as $\delta^2_{u_0}$ = 0.0644.

Table 4.3. 4: Random intercept-only model of the two levels on female genital mutilation of daughters

Fixed Effect	**Estimate**	**se**	**Exp(β)**	**95%CI**	**p-Value**
$\boldsymbol{\beta_0}$(constant)	-0.93382	0.0787	0.3930	(-1.08,-0.77)	<2e-16
Random effect					
$\boldsymbol{\delta^2_{u_0}}$ (between regions)	0.0644	0.253			
AIC = 18473.1		BIC = 18488.4	logLik = -9234.5 deviance= 18469.1		

4.3.5. Random Intercept Model for Daughters Genital Mutilation

After we realize the non-zero of variance between regions, we identify whether adding explanatory variable were needed or empty model is sufficient. The comparative statistics of the random intercept-only model (empty model) versus the random intercept model are shown in Table 4.3.5. The random intercept model with fixed explanatory variables is a better fit as compared to the empty model, because of the AIC of random intercept model (18407.9) is less as compared to that AIC of random intercept only model (18473.1). Additionally, the deviance of the empty model (18469.1) is greater than the deviance of the random intercept model (18361.9), the model with the smallest deviance is a best fitted model.

Table 4.3. 5: Comparisons of the random intercept-only model and random intercept models regarding daughter's circumcision.

	Random-intercept model	**Null(empty) model**
AIC	18407.9	18473.1
BIC	18584.1	18488.4
logLik	-9181.0	-9234.5
Deviance	18361.9	18469.1

Backward selection is used to identify the significant variables from the random intercept model. Regarding that type of residence, mother education level, father education level, religion, wealth index and media exposure variable were not significant. Age of mother and the circumcision status of the mother have a significant relation to daughter's genital mutilation/cutting.

The likelihood ratio statistic for testing the null hypothesis, that $\delta_{u_0}^2 = 0$, can be calculated by -2*(logLik(single level random model)-logLik(two level random model)) is 152.5573 (df=8) which is large and reject the null hypothesis at 0.05 level of significance. This indicates that, the variance between regions is non zero regarding daughters case of female genital mutilation. The random intercept result presented in table 4.3.6 below.

Controlling for the effects of other variables and allowing the intercept parameter to vary across regions, the likelihood of circumcised daughters from mother of age groups 20-24 and 25-29 were 13.1 percent and 18.8 percent times lower than as compared to circumcised daughters from mother of age groups 15-19, respectively. The estimated odds of genital mutilation for daughters among their mother 44-49 age groups is 1.18 times the estimated odds for daughters among their mother 15-19 age group. Finally the estimated odds of circumcision for daughters among their mother were circumcised are 1.33 times the estimated odds of circumcision for daughters among their mothers were circumcised.

Table 4.3. 6: Results of the logistic multilevel model with a random intercept for the daughters characteristics on female genital mutilation.

Fixed effect		**Estimate**	**Se**	**Exp(β)**	**95%CI**	**p-value**
	β_0(constant)	-1.0945*	0.0876	0.3346	(-1.26,-0.92)	< 2e-16
Age	15-19(ref)			1		
	20-24	-0.1393*	0.0569	0.8699	(-0.25,-0.02)	0.014363
	25-29	-0.2079*	0.0577	0.8122	(-0.32,-0.09)	0.000317
	30-34	-0.0424	0.0606	0.9584	(-0.16,0.07)	0.484453
	35-39	0.0422	0.0630	1.0431	(-0.08,0.16)	0.502468
	40-44	0.1008	0.0712	1.1061	(-0.03,0.24)	0.156930
	45-49	0.1673*	0.0784	1.1821	(0.01,0.32)	0.032948
Mother circumcision	Not circumcised			1		

status	(ref)					
	Circumcised	0.2875*	0.0408	1.3331	(0.20,0.36)	1.87e-12
Random effects						
Level(regional effects)	$\delta^2_{u_0}$	0.0602				
AIC = 18395.6	BIC = 18464.6	logLik = -9188.8	deviance = 18377.6			
Ho: $\boldsymbol{\delta^2_{u_0}}$ = 0,	$\boldsymbol{X^2}$= 152.55	logLik(single level) = -9265.075				

Note: ref = Reference categories. * Significant (P-value <0.05)

The variance components model which we have just specified and estimated above assumes that the only variation between regions is in their intercepts. We should allow for the possibility that the regions have different slopes. It is essential to determine whether the explanatory variables included in the study have a different influence on the response variable (FGM) among regions. We fitted a multilevel model with a random coefficient.

Estimates of the random coefficient model show that the random slope variances of all included variables are zero. The AIC of the random intercept model (18550) is small when compared to both random intercept-only model (18840) and random coefficient models (18554 and 18556). The model which has small AIC is the best model for the data set of FGM in Ethiopia. Additionally, the random coefficient model is not significant and the Random intercept model significantly describes the association of FGM and considered explanatory variables. The test result of Random Intercept only model, Random Intercept model, and Random Coefficient model is presented in Table 4.3.7. This indicates that the effect of these variables is the same for each region. Therefore the random intercept model is enough to analyze. Because the R software didn't allow all variables as random slopes at the same time, we include two at once.

Table 4.3.7: Test and compare three models (empty, Intercept and Coefficient models)

	Df	**AIC**	**BIC**	**logLik**	**deviance**	**Chisq**	**Chi Df**	**Pr(>Chisq)**
fit00(empty mod.)	2	18840	18855	-9418	18836			
fit0(Intercept mod.)	17	18550	18680	-9258	18516	320	15	<2e-16
fit0c1(coef. mod.)	19	18554	18699	-9258	18516	0	2	1
fit0c(coef mod.)	20	18556	18709	-9258	18516	0	1	1

4.3.6. The goodness of Fit Test

The overall model evaluation was assessed using AIC and BIC. The Deviance based chi-square test (X^2 = -2*(logLik(fitted model of less parameter)-logLik(fitted model of more parameter))) also done. Based on the result we obtained in Table 4.3.3 (random intercept model), the value of AIC and BIC are less than the model we obtained in the random intercept-only model and the deviance chi-square is significant. For daughters result in Table 4.3.6 is also similar. So, we conclude that the random intercept model is a good fit.

4.3.7. Multilevel Model Diagnostic

The diagnostic plot for residuals like Pearson and standardized residuals of the multilevel model presented under Appendix C is similar to the GEE model. In addition that from the plots of cook‘s distance under appendix C we observe that there is no any observation had cook‘s distance more than 0.35 which implies there are no effects of influential observations. The Cook’s distance less than unity showed each observation had no impact on the group of regression coefficients. A value of the leverage statistic show’s that no observation is far apart from the others in terms of the levels of the independent variables. Thus we can say that both models were not distorted.

4.4. Comparison of the Results of Multilevel and GEE

Direct comparison of multilevel and generalized estimating equations is impossible. Because their estimation method is different. In multilevel, likelihood-based methods are available for testing fit, comparing models and conducting inferences about parameters, whereas likelihood-based methods are not available in GEE to perform testing fit, comparing model.

Marginal (Population average) models estimated by GEE estimation, the estimated slope is the effect of a change in the whole population if everyone’s predictor variable changed by one unit. While on the contrary Conditional cluster specific, mixed or multilevel: the slope is the effect of change for a particular individual or subject of a unit change in a predictor.

GEE estimation is robust to assumptions about the higher level, between-group distribution; gives correct SE for the fixed part and gives the Population average value. But it does not give the higher-level variance, it is not extendible to random slopes and you cannot derive the cluster-specific estimates. All things that you get straightforwardly with the mixed, multilevel random-effects approach.

The multilevel logistic model and GEE result provided almost relatively the same value of parameters (variables coefficient). The significances of variables in both models were not contradictory to explain the considered response variables. But when we compare the standard errors of the specific coefficients of the two models, the multilevel logistic models parameters standard error were less. Thus we used multilevel logistic regression to conclude the results of this study.

4.5. Discussion

The purpose of this paper was to examine the regional differences in women and daughters genital mutilation in Ethiopia using generalized estimating equation and multilevel logistic regression procedure. The study uses the 2016 demographic and health survey data to identify some of the factors that are related to female genital mutilation. This study found evidence that some of the demographic and socioeconomic variables considered have significant relation with female genital mutilation. The region, place of residence, age, educational level and religion were found to be important factors related to female genital mutilation among reproductive-age women (15-49) years. And also the region, the age of mother and circumcised status of the mother were found to be important factors related with female genital mutilation among daughters. The results, obtained from descriptive analysis, generalized estimating equation and multi-level logistic regression have been discussed as follows.

According to the results, female genital mutilation/cutting is practiced significantly in all regions. Somali, Afar, Harari and Dire Dawa regions presented with the highest female genital mutilation among women and the lowest circumcision level observed in Gambela and Tigray regions. In all analysis of daughters' experience of FGM showed a statistically significant variation among regions in Ethiopia. Afar, Somali, Amahara, Ben-Gumuz and SNNP regions shows the highest risk of FGM practice among daughters.

In those results, age is factored positively associated with women's of FGM among reproductive-age women, that is, the experience of genital mutilation becomes higher with an increase in women's age. This finding is consistent with the result of an article of [5,1]. In our findings, the daughters from the women of age 20-24 and 25-29 were less compared to daughters from women age 15-19. Women in the higher age category (40-49) have increased odds of having daughters who experience FGM compared to 15-19 age category.

The results also shows, women with secondary and higher education were less likely to be genitally mutilated compared to women with no education. With regard to mother level of education-related with FGM of daughters, the results were unexpected. The result shows education is not significant factors of FGM among daughters. But many studies showed that, as it is a factor negatively associated with the practice of genital mutilation of daughters. [15] Reported that, daughters of mothers who are more highly educated are less likely to have undergone FGM/C than daughters of mothers with little or no education.

The finding of the study also showed that the place of residence is a significant factor contributing to having experience of FGM among reproductive age women. Rural women are mostly influenced by traditional practices than women in urban areas[12]. The results of this study also show that there is a difference in the female genital status of women residing in rural and urban areas. Women who reside in rural areas are more likely to be circumcised than women who reside in urban areas. But, place of residence is not a significant factor for FGM of daughters.

Similarly, religion has been found a significant variable for female genital mutilation. In literature, [13] said that, the women are assured the practice has religious justification and therefore they had to do it if they wanted to be Muslim women. And [13] also said that The relationship between Muslim religion and FGM practice and support found in high clustered regions such as Afar, Somali and Dire-Dawa might be due to the fact that these regions are majorly populated by Muslims and the Muslim community perceives that FGM ("Sunna type") is an important tradition and religious requirement for Muslim women. Additionally, they concluded that Muslim women had 3 times higher odds of having experienced FGM as compared to Orthodox women based on 2005 EDHS data. Our finding is consistent with these results, that is, Muslim women had 1.6 times higher odds of having experienced FGM as compared to Orthodox women. Religion is not a significant factor of FGM among daughters less than 15 years age. But [13] at that time approved, daughters from Muslim women have increased odds of experiencing FGM. This difference is may be the commitment of government regarding eliminate the traditional practice which violates human rights.

Daughters from circumcised women have increased odds of experiencing FGM compared to daughters from not circumcised women.

Wealth status has been found to be a factor related to female genital mutilation in previous studies. Higher order wealth index is a factor associated with increased FGM risk[16] and [34]. In our study wealth status is not a factor associated with FGM.

In GEE analysis, women and daughters in regions are considered as correlated and they in separate clusters are independent. Before analysis of data using GEE working correlation and standard error (empirical and model-based) are compared to use the appropriate estimation equation.

In the multilevel analysis, women among reproductive age are considering as nested within the various regions in Ethiopia. There are three multilevel models: empty model, random intercept and random coefficient model were fitted in order to explain regional differences in the women and daughters FGM. But the random coefficient model is not a significant model to explain the relation of response and covariates.

Before the analysis of data using the multilevel approach, the heterogeneity of the FGM practice with regard to regions was checked. Following this, a model without explanatory variables and two multilevel logistic regression models were fitted for the national sample as a whole in multilevel analysis.

CHAPTER FIVE

5. CONCLUSION AND RECOMMENDATION

5.1. Conclusion

This study found evidence that verify some demographic and socioeconomic variables considered in this study have a significant relation with the female genital mutilation.

Based on the findings in the preceding chapter, this study arrives at the following conclusions. Place of residence, age, educational level and religion were determined factors related to female genital mutilation among reproductive age women (15-49 years). From this study, we also conclude that Ages of mother and circumcision status of the mother were factors influence female genital mutilation among daughters. When compared these findings to the literature and previous research of the same problem, female genital mutilation is decreasing. This decline may come due to the commitment of government regarding eliminate the traditional practice which violates human rights.

From the result we conclude that the multilevel logistic model provided interesting relationships that give more precise than generalized estimating equations. The results of multilevel logistic regression analysis among all the three models, the random intercept multilevel model provided the best fit for the data under consideration. It showed that there is a significant variation in female genital mutilation between regions. Furthermore, in empty with random intercept model and random intercept models the overall variance of the constant term was found to be significant, which reflects the existence of differences in frequency of female genital mutilation across the region.

5.2. Recommendation

Based on the results the following recommendations can be made:

1. Proceeding (enforcement) of clear national policies in an active manner for the abolishment of female circumcision
2. There has to be a strong advocacy and multi-sector collaboration to stop FGM through the country.
3. Religious leaders, key informants and other stake-holders should spread the understanding that religion does not demand female genital cutting.

4. The government should give more attention to those regions with high rates of female genital mutilation. Additionally, further research on socio-cultural practices, harmful traditional practices, and other related factors should be emphasized. Actually, the prevalence of female genital mutilation is decreasing among women and daughters, but not enough yet.
5. The future studies should focus on a review of the assumptions behind the estimates and inference provided by these two models (GEE and Multilevel).
6. Urge and motivate researcher through training and education to handle missing in data on educational research to use imputation rather than deletion, as they can be implemented on the required real researches.

REFERENCE

[1] I. P. P. F. (IPPF), "IMAP Statement on the elimination of female genital mutilation," no. IMAP STATEMENT, 2015.

[2] World Health Organization, "management of health complications from female genital mutilation WHO guidelines on the," p. 64, 2016.

[3] WHO, "Female Genital Mutilation: A Student 's Manual," *WHO-International/Reproductive Heal.*, vol. WHO/RHR/01, pp. 1–104, 2001.

[4] World Health Organization, "Female genital mutilation—new knowledge spurs optimism," *Prog. Sex. Reprod. Heal. Res.*, no. 72, 2006.

[5] C. Garcia-Moreno, A. Guedes, and W. Knerr, "Understanding and addressing violence against women Female genital mutilation," vol. 1, pp. 1–8, 2012.

[6] Rose Ansorge, "Female Genital Mutilation / Cutting in the shanty towns of Port Sudan," no. February, 2008.

[7] UNICEF, *Changing a Harmful Social Convention: Female Genital Mutilation/Cutting*. 2013.

[8] Gender, H. Rights, C. Branch, and T. Division, "Global Consultation on Female Genital Mutilation/Cutting," *Tech. Rep.*, 2008.

[9] B. Vissandjée, S. Denetto, P. Migliardi, and J. Proctor, "Female genital cutting (FGC) and the ethics of care: Community engagement and cultural sensitivity at the interface of migration experiences," *BMC Int. Health Hum. Rights*, vol. 14, no. 1, 2014.

[10] 28 Too Many, *Female Genital Mutilation in Ethiopia*. 2013.

[11] Ethiopia DHS, "Ethiopia Demographic and Health Survey," *Introd. Methodol.*, 2016.

[12] K. R. Kerubo, *Female Genital Mutilation- Effects on Women and young Girls*. 2010.

[13] T. Setegn, Y. Lakew, and K. Deribe, *Geographic Variation and Factors Associated with Female Genital Mutilation among Reproductive Age Women in Ethiopia : A National Population Based Survey*. 2016.

[14] S. Daniel, S. Denetto, and L. Plenert, "Our Selves , Our Daughters : Community-Based

Education and Engagement Addressing Female Genital Cutting (FGC) with Refugee and Immigrant African Women in Winnipeg," no. April, 2011.

[15] K. Gebremariam, D. Assefa, and F. Weldegebreal, "Prevalence and associated factors of female genital mutilation among Somali refugees in eastern Ethiopia: a cross-sectional study," vol. 9, no. 264, pp. 357–365, 2016.

[16] R. Aseffa, *Factors Affecting the practice of Female Genital Mutilation of Ethiopian Women*. 2011.

[17] G. Dahmen and A. Ziegler, "Generalized Estimating Equations in Controlled Clinical Trials : Hypotheses Testing," vol. 46, pp. 214–232, 2004.

[18] P. Ghisletta, "An Introduction to Generalized Estimating Equations and an Application to Assess Selectivity Effects in a Longitudinal Study on Very Old Individuals," vol. 29, no. 4, pp. 421–437, 2015.

[19] K.-Y. Liang and Scott L . Zeger, "Longitudinal Data Analysis Using Generalized Linear Models," vol. 73, no. 1, pp. 13–22, 1986.

[20] H. X. Barnhart and J. M. Williamson, "Goodness-of-fit tests for GEE modeling with binary responses.," *Biometrics*, vol. 54, no. 2, pp. 720–729, 1998.

[21] W. Tang, N. Lu, R. Chen, and H. Zhang, "Longitudinal Data Analysis," 2013.

[22] H. Goldstein, *Multilevel Statistical Models*. 1999.

[23] R. S. Stawski, "A Multilevel Analysis : An Introduction to Basic and Advanced Multilevel Modeling (2nd Edition)," no. October 2014, pp. 37–41, 2013.

[24] G. D. Garson, *Fundamentals of Hierarchical Linear and Multilevel Modeling*. 2011.

[25] J. A. N. F. P. McCullagh, *Generalized Linear Model*, Second Edi. London and new York, 1983.

[26] B. Winter, "linear mixed effects analyses," no. Tutorial 2, pp. 1–22, 2014.

[27] Central Statistical Agency, "Federal Democratic Republic Of Ethiopia Demographic and Health Survey," 2016.

[28] A. Agresti, *Categorical Data Analysis*, Second Edi. Gainesville, Florida, 2002.

[29] G. Molenberghs and G. Verbeke, *Models for Discrete Longitudinal Data*, vol. 53, no. 9. 2003.

[30] I. Owusu-Darko, I. K. Adu, and N. K. Frempong, “Application of generalized estimating equation (GEE) model on students’ academic performance,” *Appl. Math. Sci.*, vol. 8, no. 68, pp. 3359–3374, 2014.

[31] G. A. Ballinger, “Using Generalized Estimating Equations for Longitudinal Data Analysis,” *Organ. Res. Methods*, vol. 7, no. 2, pp. 127–150, 2004.

[32] K.-Y. Liang, Scott, and S. L. Zeger, *Analysis of Longitudinal Data*. 2002.

[33] V. J. Carey and Y. Wang, “Working covariance model selection for generalized estimating equations,” no. August 2008, 2011.

[34] J. M. Gran, L. Wasmuth, E. J. Amundsen, B. H. Lindqvist, and O. O. Aalen, “Growth rates in epidemic models: application to a model for HIV/AIDS progression,” *Stat. Med.*, vol. 28, no. July 2006, pp. 221–239, 2009.

[35] M. Nakai and W. Ke, “Review of the Methods for Handling Missing Data in Longitudinal Data Analysis,” vol. 5, no. 1, pp. 1–13, 2011.

[36] C. Szmaragd and G. Leckie, “Multilevel Models for Binary Responses R Practical Introduction,” *Matrix*, vol. 7, 2011.

[37] M. J. Jang, “Working correlation selection in generalized estimating equations,” 2011.

APPENDIXES

Appendix A

Descriptive

```
rm(list=ls()) ###### to clear the work space

library(foreign)

genitaldata <- read.spss( "C:\\Users\\Des\\Desktop\\ETIR70FL.sav", to.data.frame=T, use.value.labels = F, head=T) ### to read data from spss

gen.data<-genitaldata[,c(1,16,45,46,51,66,102,4864,4866,4915,4935)]

###########to variable name#############

names(gen.data)[2:11] <- c("agec","region", "residt", "educs", "religion", "wealthi", "G100", "w.fgms" , "d.gms", "d.agec")

####################################for women data missing imputation

library(Hmisc)

wfgms.imp<- impute(gen.data$w.fgms, "random"); x<-as.data.frame(wfgms.imp)

cc<-cbind(gen.data, wfgms.imp)

##################################################for daughters data missing imputation

dgms.imp<- impute(cc$d.gms, "random"); xy<-as.data.frame(dgms.imp)

ccd<-cbind(cc, dgms.imp)

pp1 <- prop.table(table(ccd$w.fgm)) * 100

barplot(pp1 ,xlab="women circumcised=1 or not=0", ylab="percentage(%)", main="The circumcision status of women in 2016")

pp2<- prop.table(table(ccd$d.gms)) * 100

barplot(pp2,xlab="daughters circumcised=1 or not=0", ylab="percentage(%)", main="The circumcision status of daughters in 2016")

library(gmodels) ###to calculate crosstable

library(gdata); library(gtools)

CrossTable(ccd$w.fgms,ccd$region,digits=2, expected=TRUE,dnn=c("w.fgms","region"))

CrossTable(ccd$w.fgms,ccd$region,digits=2, chisq=TRUE, dnn=c("w.fgms","region"))
```

```
CrossTable(ccd$w.fgms,ccd$region, format=c("SPSS"), digits=1)#############the best way of calculating prevalence for women

CrossTable(ccd$w.fgms,ccd$residt, format=c("SPSS"), digits=1); CrossTable(ccd$w.fgms,ccd$agec, format=c("SPSS"), digits=1)

CrossTable(ccd$w.fgms,ccd$educs, format=c("SPSS"), digits=1); CrossTable(ccd$w.fgms,ccd$wealthi, format=c("SPSS"), digits=1)

CrossTable(ccd$w.fgms,ccd$religion, format=c("SPSS"), digits=1)

CrossTable(ccd$d.gms,ccd$region, format=c("SPSS"), digits=1)## prev. for daugh

CrossTable(ccd$d.gms,ccd$residt, format=c("SPSS"), digits=1); CrossTable(ccd$d.gms,ccd$agec, format=c("SPSS"), digits=1)

CrossTable(ccd$d.gms,ccd$educs, format=c("SPSS"), digits=1); CrossTable(ccd$d.gms,ccd$wealthi, format=c("SPSS"), digits=1)

CrossTable(ccd$d.gms,ccd$religion, format=c("SPSS"), digits=1)
```

```
meds<- table(ccd$w.fgms,ccd$region)/sum(table(ccd$region)) * 100

bp<- barplot(meds, beside=TRUE, axes=FALSE, xlab="circumcision status of women",

names=c("Tig.", "Afar","Amah.", "Orom.", "Soma.", "Beni", "SNNP", "Gamb.", "Hara.","A.ababa","D.dawa"), col=c("azure3", "azure"),

ylab="Frequency (%)", ylim=c(0,5))

axis(2, at=seq(0,10,2))

legend("topright", legend=c("No", "Yes"), bty="n", fill=c("azure3", "azure"))

text(bp, 0, round(meds, 1), cex=1, pos=3)

############################################3##3to select model using QIC#####

##############################this is workingQIC CIC

QIC.binom.geeglm<-  function(model.geeglm, model.independence)

{

AIinverse<- solve(model.independence$geese$vbeta.naiv)

V.msR<- model.geeglm$geese$vbeta

trace.term<- sum(diag(AIinverse%*%V.msR))

mu.R<- model.geeglm$fitted.values
```

y <- model.geeglm$y

scale<- 1

quasi.R<- sum(y*log(mu.R/(1-mu.R))+log(1-mu.R))/scale

QIC <- (-2)*quasi.R + 2*trace.term

output<- c(QIC,trace.term)

names(output) <- c('QIC','CIC')

output

}

sapply(list(ftwex1,ftwex2,ftwind1,ftwind2,ftwar11,ftwar12),

function(x) QIC.binom.geeglm(x,c(ftwex1,ftwex2,ftwind1,ftwind2,ftwar11,ftwar12)))

#######to fit GEE model########

library(geepack)

ftwex1<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+ factor(religio) +factor(wealth), data=geni, family= binomial("logit"), corstr = "exchangeable", id=regio, std.err="san.se")

summary(ftwex1)#####without impute and id=region; anova(ftwex1)

ftwex2<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+ factor(religio) +factor(wealth), data=geni, family= binomial("logit"),corstr = "exchangeable", id=regio, std.err="fij")

ftwind1<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio) +factor(wealth), data=geni, family= binomial("logit"),corstr = "independence", id=regio, std.err="san.se")

ftwind2<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio) +factor(wealth),, data=geni, family= binomial ("logit"), corstr = "independence", id=regio, std.err="fij")

ftwar11<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio) +factor(wealth), data=geni, family= binomial("logit"),corstr = "ar1", id=regio, std.err="san.se")

summary(ftwar11)

ftwar12<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio) +factor(wealth), data=geni, family= binomial("logit"),corstr = "ar1", id=regio, std.err="fij")

###############t0 fit Gee model after correlation structure was selected and not significant vari. ommitted

ftwex1o<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+ factor(religio), data=geni, family= binomial("logit"), corstr = "exchangeable", id=regio, std.err="san.se")

summary(ftwex1o)#####without impute and id=region; anova(ftwex1)

ftwex2o<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+ factor(religio), data=geni, family= binomial("logit"), corstr = "exchangeable", id=regio, std.err="fij")

ftwind1o<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio), data=geni, family= binomial("logit"), corstr = "independence", id=regio, std.err="san.se")

ftwind2o<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio) , data=geni, family= binomial ("logit"), corstr = "independence", id=regio, std.err="fij")

ftwar11o<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio), data=geni, family= binomial("logit"), corstr = "ar1", id=regio, std.err="san.se")

summary(ftwar11o)

round(summary(ftwar12o)$coefficients,2)######have have limited decimal

ftwar12o<-geeglm(fgms.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio), data=geni, family= binomial("logit"), corstr = "ar1", id=regio, std.err="fij")

round(summary(ftwar12o)$coefficients,2)######have have limited decimal

###to calculate QIC based on the above written function

sapply(list(ftwex1,ftwex2,ftwind1,ftwind2,ftwar11,ftwar12),

function(x) QIC.binom.geeglm(x,c(ftwex1,ftwex2,ftwind1,ftwind2,ftwar11,ftwar12)))

###to fit model for daughters####################################

ftd<-geeglm(dgm.imp~factor(age)+ factor(regio)+factor(resid)+factor(educ)+factor(religio)+factor(wealth) +factor(mediaexposure) +factor(fgms.imp)+factor(fatherseduc), data=geni,family=binomial("logit"), corstr = "ar1", id=regio,scale.fix=T,std.err="san.se") summary(ftd); anova(ftd)fit the model of daughters case

ftd100<-geeglm(dgms.imp~factor(region)+factor(wfgms.imp),
data=geni,family=binomial("logit"), corstr = "ar1", id=region, scale.fix=T,std.err="san.se")

summary(ftd100); round(summary(ftd100)$coefficients,3); anova(ftd100)

ftd101<-geeglm(dgms.imp~factor(region)+factor(wfgms.imp),data=geni,
family=binomial("logit"), corstr="ar1", id=region, scale.fix =T, std.err ="fij")

summary(ftd101);anova(ftd101); exp(summary(ftd101)$coefficients)

################

ftd100<-geeglm(dgms.imp~ factor(region)+factor(wfgms.imp), data=geni,family= binomial("logit"), corstr = "ar1", id=region,scale.fix=T,std.err="san.se")

summary(ftd100); round(summary(ftd100)$coefficients,3); anova(ftd100)

ftdoo<-geeglm(dgm.imp~factor(age)+ factor(regio)+ factor(fgms.imp), data= geni, family= binomial("logit"),corstr = "ar1", id=regio,scale.fix=T,std.err="san.se")

summary(ftdoo); anova(ftdoo); round(summary(ftdoo)$coefficients,3)

ftdoo1<-geeglm(dgm.imp~ factor(age)+ factor(regio)+ factor(fgms.imp), data= geni, family= binomial("logit"),corstr = "independence", id=regio,scale.fix=T,std.err="san.se")

summary(ftdoo1); round(summary(ftdoo1)$coefficients,3)

ftdoo2<-geeglm(dgm.imp~factor(age) + factor(regio)+ factor(fgms.imp), data= geni, family= binomial("logit"),corstr = "exchangeable", id=regio,scale.fix=T,std.err="san.se");

summary(ftdoo2); round(summary(ftdoo2)$coefficients,3)

ftdoo3<-geeglm(dgm.imp~factor(age)+ factor(regio)+ factor(fgms.imp), data=geni, family=binomial("logit"), corstr = "ar1", id=regio, scale.fix=T,std.err="fij")

summary(ftdoo3); round(summary(ftdoo3)$coefficients,3)

ftdoo4<-geeglm(dgm.imp~factor(age)+ factor(regio) + factor(fgms.imp), data=geni, family=binomial("logit"), corstr = "independence", id=regio, scale.fix=T,std.err="fij")

summary(ftdoo4); round(summary(ftdoo4)$coefficients,3)

ftdoo5<-geeglm(dgm.imp~factor(age)+ factor(regio) +factor(fgms.imp), data=geni, family= binomial("logit"), corstr = "ar1", id=regio, scale.fix=T,std.err="fij")

summary(ftdoo5); round(summary(ftdoo5)$coefficients,3)

sapply(list(ftdoo, ftd101, ftdoo2, ftdoo3, ftdoo4, ftdoo5**),**

The Multilevel r code

library(lme4); library(nloptr)

fit00 <- glmer(fgms.imp ~ 1+(1|regio), family = binomial("logit"), data = geni, nAGQ = 5);summary(fit00)##null two-level model

fit000<-glm(fgms.imp ~ 1,family = binomial("logit"), data=geni)###null single level model

logLik(fit000)

-2*(logLik(fit000)-logLik(fit00))####

###$$$$t o create 95%CI

se<-sqrt(diag(vcov(fit00))); (tab <- cbind(Est = fixef(fit00), LL = fixef(fit00) - 1.96 * se, UL = fixef(fit00) + 1.96 *se))

##to plot the constant of random effect(regional variability)

u0 <- ranef(fit00, condVar = TRUE) ; u0se <- sqrt(attr(u0[[1]], "postVar")[1, ,])

region<- as.numeric(rownames(u0[[1]])); u0tab <- cbind("region" = region, "u0" = u0[[1]], "u0se" = u0se)

colnames(u0tab)[2] <- "u0"; u0tab <- u0tab[order(u0tab$u0),] ; u0tab <- cbind(u0tab, c(1:dim(u0tab)[1]))

u0tab <- u0tab[order(u0tab$region),] ; colnames(u0tab)[4] <- "u0rank"

plot(u0tab$u0rank, u0tab$u0, type = "n", xlab = "u_rank", ylab = "conditional

modes of r.e. for region:_cons", ylim = c(-1, 1), main="The plot of the estimated residuals for for all regions")

segments(u0tab$u0rank, u0tab$u0 - 1.96*u0tab$u0se, u0tab$u0rank, u0tab$u0 + 1.96*u0tab$u0se)

points(u0tab$u0rank, u0tab$u0, col = "blue"); abline(h = 0, col = "red")

####to compare random intercept only model and random intercept model

fit0 <- glmer(fgms.imp ~ factor(resid)+factor(age)+factor(educ)+factor(religio)+(1| regio), family = binomial("logit"),

data = geni, nAGQ = 5);summary(fit0)#######using adaptive gauss-hermite fit011 <- glm(fgms.imp ~ factor(resid)+factor(age)+factor(educ)+factor(religio), family = binomial("logit"), data = geni); summary(fit011)

-2*(logLik(fit011)-logLik(fit0)); anova(fit0)

se<-sqrt(diag(vcov(fit0))); (tab <- cbind(Est = fixef(fit0), LL = fixef(fit0) - 1.96 * se, UL = fixef(fit0) + 1.96 *se))

###to chekglmer with glmmPQL(maximum likelihood using AGHQ and PQL)

###to test model

anova(fit0 ,fit00, test="Chi")###with chi-sqr ###wow #wowo #wowo wow

fi <- glmmPQL(fgms.imp ~ factor(age)+ factor(resid)+ factor(educ)+ factor(religio)+ factor(wealth), (~1|regio), family = binomial("logit"), data = geni) ; summary(fi)

daughters case

fitd0 <- glmer(dgm.imp ~ factor(age)+ factor(fgms.imp)+ (1| regio),family = binomial("logit"), data = geni , nAGQ = 5); summary(fitd0); exp(summary(fitd0)$coefficients)

se<-sqrt(diag(vcov(fitd011))); (tab <- cbind(Est = fixef(fitd011), LL = fixef(fitd011) - 1.96 * se, UL = fixef(fitd011) + 1.96 *se))

fitd02 <- glm(dgm.imp ~ factor(age)+factor(fgms.imp),family = binomial("logit"), data = geni); summary(fitd02)

-2*(logLik(fitd02)-logLik(fitd0))

fid <- glmmPQL(dgm.imp ~ factor(age) + factor(resid)+factor(educ) + factor(religio)+ factor(wealth)+ factor(mediaexposure) + factor(fgms.imp), (~1|regio), family = binomial("logit"), data = geni); summary(fid)

############coefficient Model

fit0c <- glmer(fgms.imp ~ factor(resid)+factor(age)+factor(educ)+factor(religio)+(1|regio)+(1|resid)+(1|age), family = binomial("logit"), data = geni, nAGQ = 1)

summary(fit0c)

fit0c1 <- glmer(fgms.imp ~ factor(resid)+factor(age)+factor(educ)+factor(religio)+(1|regio)+(1|educ)+(1|religio), family = binomial("logit"), data = geni, nAGQ = 1)

summary(fit0c1)

anova(fit00,fit0,fit0c1 ,fit0c, test="Chi")###with chi-sqr

############Diagnostic checking for GEE models

plot(resid(ftwar11))

plot(fitted(ftwar11),resid(ftwar11))

Appendix B

Plot*1: The Plot of the estimated residuals (random effects) for all regions (for daughter's model)

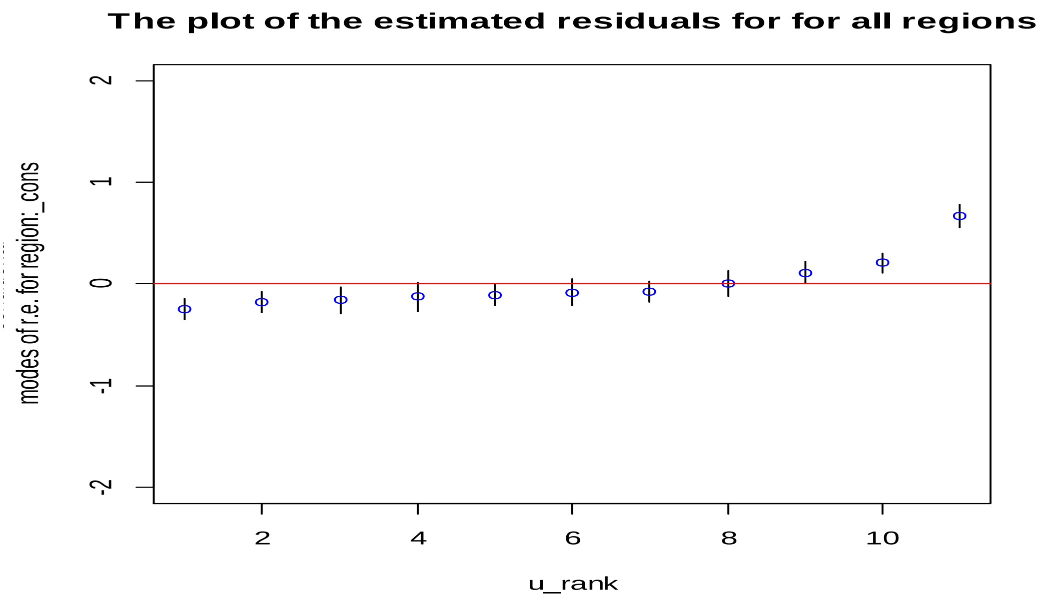


Table 1A Frequency of women (respondent) circumcised

Respondent(women) circumcised

		Frequency	Percent	% of resp.
Response	No	4099	25.8	29.6
	Yes	7286	46.8	70.4
	Total	11385	72.6	100
Missing	System	4293	27.4	
Total		**15683**	**100**	

Table 1B Frequency of daughter's circumcision

daughters circumcision

		Frequency	Percent	% of respon
Response	No	7434	53.5	77.2
	Yes	1692	10.8	22.8
	Total	10069	64.3	100
Missing	System	5614	35.7	
Total		**15683**	**100**	

Table1 2: Test for the goodness of fit and model selection from correlation structure of auto regressive, independence and exchangeable.(for case of daughters FGM)

	Autoregressive (ftd10)	**Independence (ftd101)**	**Exchangeable (ftd102)**
QIC	18378.3	18378.4	18378.9
CIC	22.1	22.1	22.4

Table 3: the null model of single level

Fixed effect	**estimate**	**Degree of freedom**
$\boldsymbol{\beta_0}$(constant)	0.8087	15682
AIC: 19370	logLik: -9686.245 (df=1)	

Table 4: Wald test of vaiables in GEE model for women FGM

	Df	X2	P(>|Chi|)
factor(age)	6	93	< 2e-16 ***
factor(regio)	10	388	< 2e-16 ***
factor(resid)	1	33	9.5e-09 ***
factor(educ)	3	51	4.1e-11 ***
factor(religio)	5	49	2.3e-09 ***
factor(wealth)	4	9	0.058 .

Table*4A: Wald test of vaiables in GEE model for daughters FGM

	Df	X2	P(>|Chi|)
factor(age)	6	36.1	2.7e-06 ***

factor(regio)	10	165.0	< 2e-16 ***
factor(resid)	1	1.9	0.17
factor(educ)	3	2.8	0.43
factor(religio)	5	4.3	0.50
factor(wealth)	4	7.8	0.10 .
factor(mediaexposure)	1	0.0	0.86
factor(fgms.imp)	1	43.6	4.0e-11 ***
factor(fthredu.imp)	4	9.5	0.06

Table*5A: anova test of variables in multilevel model for women FGM

	Df	Sum Sq	Mean Sq	F value
factor(resid)	1	63.8	63.8	63.85
factor(age)	6	101.5	16.9	16.92
factor(educ)	3	55.4	18.5	18.48
factor(religio)	5	98.2	19.6	19.64
factor(wealth)	4	11.0	2.7	2.74

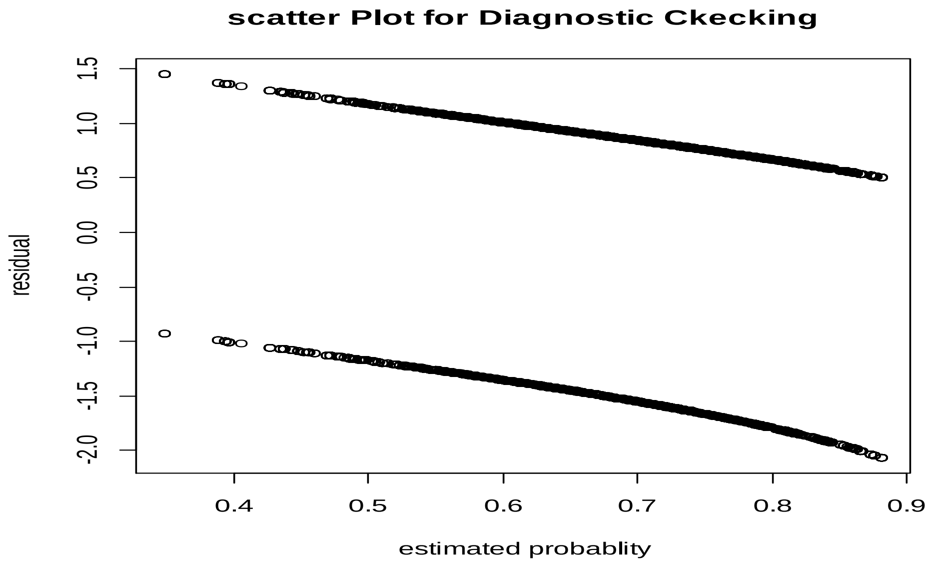
Table*5A: anova test of variables in multilevel model for Daughters FGM

	Df	Sum Sq	Mean Sq	F value
factor(age)	6	40.2	6.7	6.69
factor(resid)	1	2.9	2.9	2.89
factor(educ)	3	3.2	1.1	1.06
factor(religio)	5	4.4	0.9	0.87
factor(wealth)	4	8.0	2.0	2.01
factor(mediaexposure)	1	0.0	0.0	0.03
factor(fgms.imp)	1	46.8	46.8	46.83
factor(fthredu.imp)	4	10.5	2.6	2.03

Appendix C

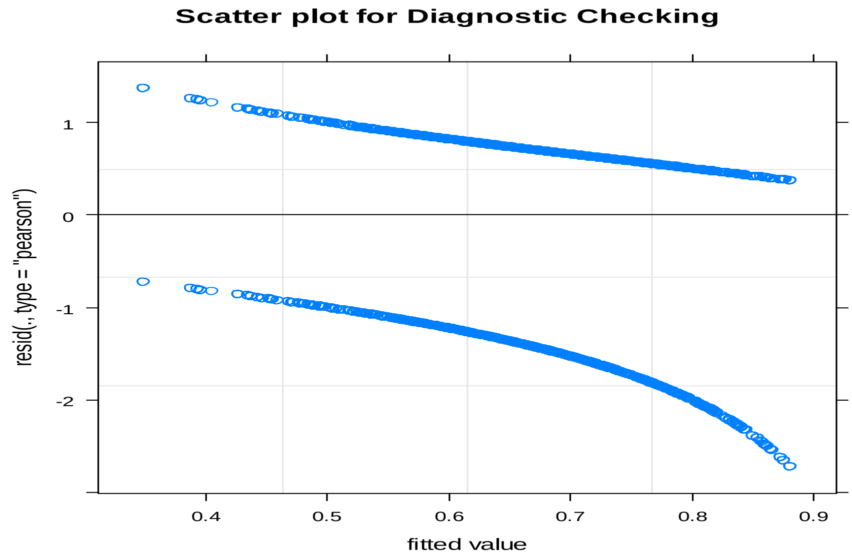
plot(fitted(fit0),resid(fit0), main="scatter Plot for Diagnostic Ckecking",xlab="estimated probablity" , ylab= “residual")

(almost the same for GEE models)



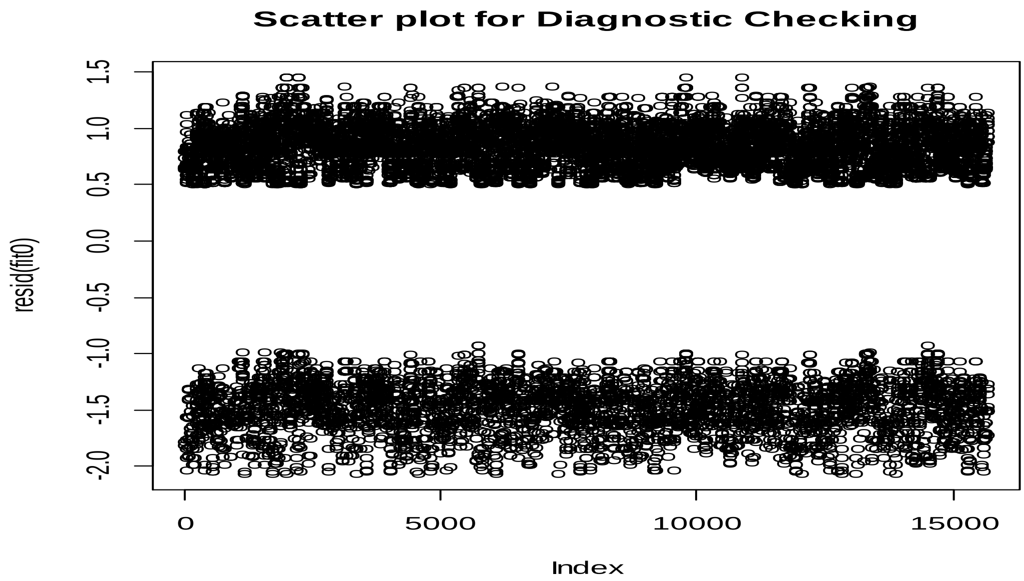
plot(fit0, main="Scatter plot for Diagnostic Checking", xlab="fitted value")

(qqnorm(residuals(fit0)))



plot(resid(fit0),main="Scatter plot for Diagnostic Checking")

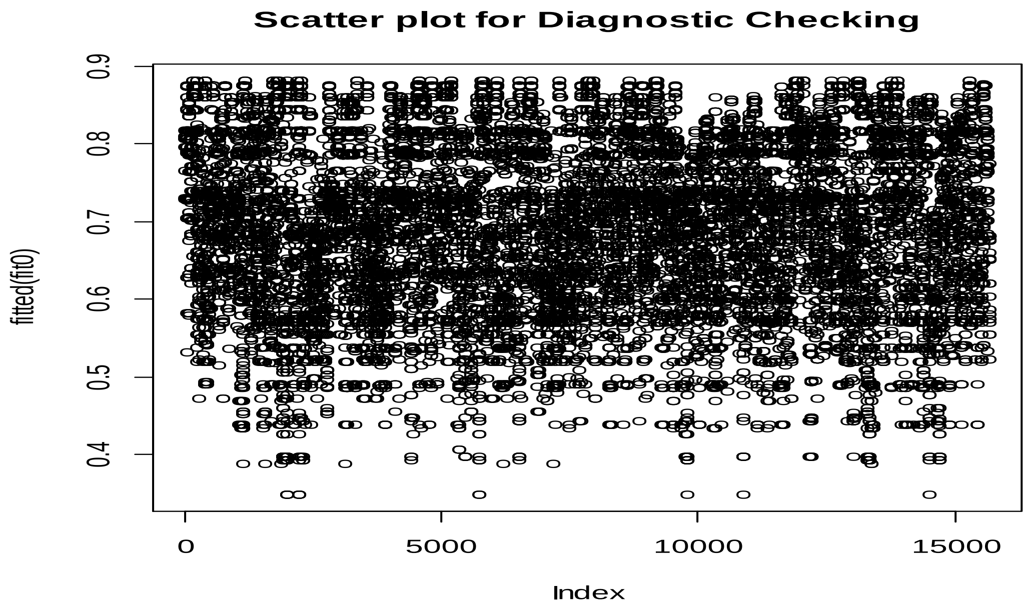
(almost the same for GEE models)



plot(fitted(fit0),main="Scatter plot for Diagnostic Checking")

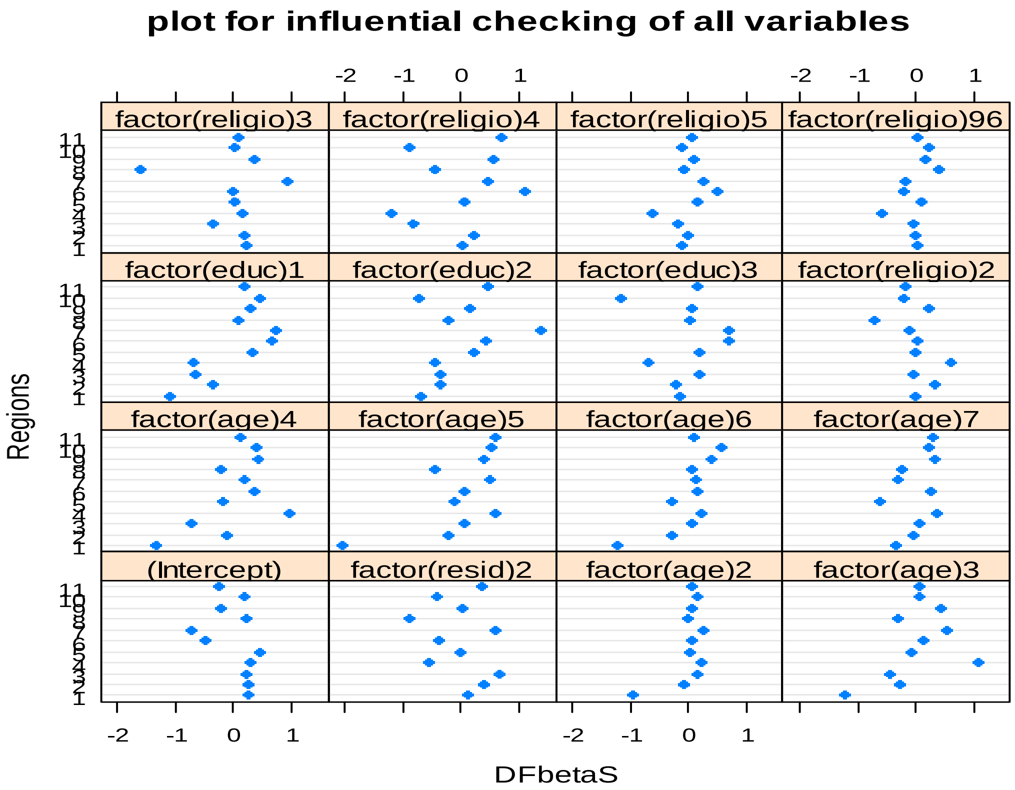
(stripplot(resid(fit0), main="strip plot for Diagnostic checking"))

(almost the same for GEE models)



plot(estex.fit0,which="dfbetas",xlab="DFbetaS", ylab="Regions",

main="plot for influential checking of all variables")



plot(cooks.distance(estex.fit0),xlab="region", main ="Cook's distance")