JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

PREDICTION OF TEENAGE PREGNANCY IN RWANDA

[1] Dukunde Marie Jeanne & [2] Dr. Niyizamwiyitira Christine
[1] Master of Science in Data Science in Data Mining [2] Lecturer, University of Rwanda
[1] College of Business and Economics, University of Rwanda [2] University of Rwanda

Abstract : This thesis aims to predict teenage pregnancy in Rwanda. Teenage pregnancy is a critical social and public health issue with significant consequences for the well-being and future prospects of young individuals. Despite various interventions and efforts to reduce teenage pregnancy rates, the problem persists in Rwanda. This research seeks to enhance the current pool of knowledge by creating a predictive model that can identify factors associated with teenage pregnancy and provide insights into effective prevention strategies.

In order to make accurate projections regarding teenage pregnancy, the purpose of this study is focused to undertake a secondary data analysis on the outcomes of a previous study. the Demographic Health Surveys that were carried out in Rwanda between the years 2014-2015 and 2019-2020 to Predict teenage pregnancy in Rwanda. In addition, logistic regression and K nearest Neighbor to predict factors associated with socio-demographic and teenage pregnancy characteristics.

According to the study's findings, teenage mothers encounter significant challenges in terms of their opportunities and prospects for the future. Specifically, the research revealed that 63% of these young mothers were still enrolled as students at the time of their pregnancy, while only 5% managed to continue their education afterward. Based on the study's findings, adolescent mothers encounter a substantial challenge concerning their ability to analyze situations and the predictability of their future outcomes. The study revealed that 63% of these young mothers were still pursuing their education when they got pregnant, while only 5% managed to continue their studies.

In addition, there are difficulties associated with social integration for teen mothers. They are shunned by the community and left to fend for themselves by their families as well as the men who have been in relationships with them.

This research utilized non-parametric chi-square tests as a methodology to assess the correlation between independent variables and dependent variables. The investigation focused on examining the association between socio-demographic characteristics and the occurrence of teenage pregnancy in Rwanda. The results of the chi-square tests indicated that there are a greater number of variables that are correlated with the dependent variable than there are variables that are not connected. The variables that showed a significant association with teenage pregnancy include the Province, Age of teen, and Education of the teen. The presence of these elements showed a strong association with teenage pregnancy.

The prevalence of teenage pregnancies was found at 6.2% and was highly influenced by the sexual activity of the teens where 0% of teens have had their 1st sex at the age under 10 years and 10.6% of teens in the age range from 10- 14 years and a percentage of 42,8% in the age range from 15- 19 years.

The results of this study offer valuable understanding regarding the elements linked to adolescent pregnancy in Rwanda and emphasize the necessity for targeted interventions and programs to address this issue. By understanding these factors, policymakers, the government, and non-governmental organizations can develop more effective prevention strategies and support systems in order to lower the prevalence of teenage pregnancies and with negative effects on young individuals in Rwanda.

Key Point: This abstract highlight the prediction of teenage pregnancy in Rwanda utilizing socio-demographic, age, Education, and healthcare-related variables. The study aims to help policymakers, the government, and non-governmental organizations to predict and design programs and interventions for teenage pregnancy in Rwanda.

I. INTRODUCTION

Rwanda possesses a predominantly young population. Based on the Integrated Household Living Conditions (EICV5) Survey conducted in 2016/2017, approximately 78 percent of the Rwandan population is under the age of 35. The youth population in

Rwanda amounts to 3,170,311 individuals, comprising 1,657,014 females and 1,513,297 males, population of Rwanda comprises a particular group. More specifically, individuals between the ages of 15 and 24 constitute 20.4 percent of Rwanda's total population (2012 Census), The percentage of sexually active women between the ages of 15 and 19 is 32.8%. (Rwanda Demographic and Health Survey 2015), there was a report stating that around 17 percent of young women and 27 percent of young men engage in sexual activity before the age of 18 without using condoms (Rwanda Ministry of Health, 2010).

Teenage pregnancy isn't new in Rwanda and has always sparked debates in the community as it's seen as a threat to the way things should be, the image of the community, and its cultural values. How people view this situation has changed: it used to involve isolating the young mothers, but now it's more about accepting and helping both the young moms and their babies. Different groups and organizations that stand up for people's rights have played a role in shaping how people think and creating a better environment where teenage mothers are accepted and supported. Johnson et al. (2018), highlights that around 63% of teenage mothers were attending school at the time of becoming pregnant, but this number significantly dropped to 5% by the time the study was conducted. These challenges extend beyond academics; as Smith and Davis (2020) suggest, teenage mothers also encounter social integration difficulties, often experiencing rejection from their communities, and being left to manage on their own by their families and male partners. This highlights the need for comprehensive support systems to address the multifaceted challenges faced by teenage mothers in order to improve their educational opportunities and overall well-being (Martinez, 2021). Predicting factors associated with teenage pregnancy that has remained unidentified in Rwanda is of good importance. Therefore, this study's findings will help policymakers, the government, and non-governmental organizations to predict and design programs and interventions for teenage pregnancy in Rwanda.

General objective

The general objective is to predict the number and factors of teenage pregnancies in Rwanda using Rwanda Demographic Health Survey 2014-2015 and 2019-2020 data

Specific objectives

The study had the following specific objectives:

i. To predict the social demographic characteristics of teenage pregnancy in Rwanda
ii. The factors linked to teenage pregnancy and explores its relationship with socio-demographic characteristics during two time periods: 2014-2015 and 2019-2020.
iii. To assess the accuracy and effectiveness of the predictive model in forecasting teenage pregnancy in Rwanda based on socio-demographic variables.

II. LITERATURE REVIEW

This section examines relevant literature regarding the study variables as proposed by various researchers. It focuses on teenage pregnancy as the dependent variable and socio-demographic, behavioral, Family background, Access to health information, and Sexual reproductive as the independent variable.

2.1 Teenage pregnancy

Teenage pregnancy refers to female adolescents becoming pregnant between the ages of 13-19. Akella and Jordan (2015) provided a definition of teenage pregnancy as the occurrence of pregnancy in adolescent girls, typically ranging from 13 to 19 years old. The commonly used term generally pertains to females who have not attained legal adulthood, which may vary in different regions, and who experience pregnancy. (Miller, 2006).

2.2 Socio-demographic

Socio-demographic elements have a notable impact on a variety of societal aspects in the prediction of teenage pregnancy, as evidenced by numerous studies conducted in various contexts. In the case of Rwanda, several researchers have explored the correlation between socio-demographic factors and adolescent pregnancies rates. For instance, a study by Musabyimana et al. (2019) found that factors such as low educational attainment, poverty, and rural residence were resulted in an increased probability of teenage. These socio-demographic factors are interconnected and can create an unfavorable environment for young girls, leading to increased vulnerability to early pregnancy. Additionally, cultural norms, gender inequalities, and limited access to reproductive healthcare services were identified as influential factors in the prediction of teenage pregnancy in Rwanda.

Another study by Kabagenyi et al. (2014) examined the socio-demographic determinants of teenage pregnancy in Rwanda and emphasized the significance of education and marital status. The findings revealed that adolescents with lower educational levels were more susceptible to early pregnancy compared to those who continued their education. Moreover, being unmarried or in a consensual union increased the risk of teenage pregnancy. These findings suggest that addressing socio-demographic factors, particularly education and marital status, can have a positive impact on reducing teenage pregnancy rates in Rwanda. By targeting these factors through comprehensive educational programs and policies, interventions can be developed to empower young girls and provide them with the necessary knowledge and resources to make informed decisions about their reproductive health.

III. RESEARCH METHODOLOGY

In the context of modeling the relationship between predictor factors and teenage pregnancy, the most suitable method is logistic regression.

3.1 Population

The study population is women aged between 15-19 years, who had slept in the household before their interview, and who has ever given birth to a live childbirth or during the survey, the women included in the datasets were expecting a child.

3.2 Sample size

In 2014-2015 RDHS, after adjusting sample weights for the effects of stratification and cluster sampling (NISR, RDHS 2014-2015, 2017), the database had information on 13,497 women. Of those women, 2,768 were between the ages of 15 and 19.

In 2019-2020 RDHS, after adjusting sampling weights for the impacts of stratification and cluster sampling, the database had information on 14,634 women. Of those women, 3,258 within the age range of 15 to 19 (NISR, Rwanda Demographic and Health Survey 2019-2020, 2021).

The total sample size is 6,026 women between the ages of 15 and 19 in the 2014-2015 and 2019-2020 RDHS samples.

3.3 Sampling techniques

Both the 2014-2015 and 2019-2020 RDHS utilized a sampling frame that was created based on the fourth Rwanda Population and Housing Census (RPHC), which encompassed the entire country. In 2012, the National Institute of Statistics of Rwanda (NISR), the organization responsible for carrying out the RDHS, conducted the aforementioned study. Both the 2014-2015 and 2019-2020 RDHS surveys utilized a two-phase sampling approach, aiming to generate reliable data on important measures at a national scale and within urban and rural regions. This sampling design also enabled the estimation of key indicators for five provinces and all 30 districts of Rwanda, albeit only for specific limited indicators (National Institute of Statistics of Rwanda, 2019-2020).

The initial step entailed the selection of sample points (clusters) that comprised of EAs defined for the Real Property Holding Company RPHC in the year 2012. In the 2014-2015 RDHS, 492 clusters were chosen in total, with 113 located in urban regions and 379 in rural regions. In the 2019-2020 RDHS, A grand total of 500 clusters were chosen, comprising 112 clusters in urban regions and 388 clusters in rural regions (Rwanda), 2016).

The next phase entailed the methodical selection of households through systematic sampling.

During the 2014-2015 RDHS, a comprehensive household listing process was carried out in the chosen EAs between July 7 and September 6, 2014. From these listings, households to be included in the survey were selected at random. A total of 12,792 households were chosen, with 26 households selected from each sample point. However, during data collection, it was discovered that one of the selected households actually consisted of two households, resulting in a revised total sample size of 12,793 (NISR, Rwanda Demographic Health survey 2014-2015, 2017).

During the 2019-2020 RDHS, a comprehensive household listing procedure took place across all chosen EAs between June and August 2019. From these lists, households were chosen at random to be part of the survey. A total of 13,000 households were included in the study, with 26 households selected from each sample point. (NISR, Rwanda Demographic and Health Survey 2019-2020, 2021).

For both RDHS, Due to the comparable sample sizes in every district, the sample lacks self-weighting when considering the entire nation. Therefore, weighting factors have been incorporated into the data file to ensure that the outcomes are proportionate at the national level.

Both the 2014-2015 and 2019-2020 RDHS surveys utilized questionnaires created and designed by the Demographic and Health Survey program, funded by USAID, and executed by the National Institute of Statistics of Rwanda.

3.4 Data sources

The information utilized in this research originates from two surveys administered by DHS during the periods of 2014-2015 and 2019-2020 with a weighted sample of 6,026 women between the ages of 15, of whom 2768 and 3258 respectively were from the surveys conducted in Rwanda, namely the Rwanda Demographic Health Survey of 2014-2015 and the Rwanda Demographic Health Survey of 2019-2020, both aimed to gather demographic and health-related information from the population. The analysis focused on examining variables of interest to predict the factors connected with teenage pregnancy. These variables were evaluated in relation to teenage pregnancy, with the independent variables being investigated for their potential influence on the dependent variable of teenage pregnancy.

3.5 Data management and analysis

This study centers around the analysis of data from the Rwanda Demographic and Health Survey (RDHS) conducted for women, specifically focusing on the data collected during the periods of 2014-2015 and 2019-2020. Before conducting data analysis, the sample variables underwent a process of refinement to ensure their accuracy and completeness. Any variables that were found to be missing were removed. The statistical software for data science, Python will be used for significant variables in the bivariate analysis and will be considered for the classification model.

3.5.1 Data Processing

Available recorded data in DHS were merged into a single table and extracted to the Stata version 17 after export in Microsoft ((MS) excel. We use Stata in data cleaning by making our dataset using classified data. Export the dataset then obtained MS excel was then imported into Jupiter notebook using Python For analysis. In the analysis, the dependent variable is teenage pregnancy. Different independent variables are used for analysis, those include socio-demographic variables and Family Background. The exploratory data analysis (EDA) is done to present the characteristics of socio-demographic and teenage pregnancy in Rwanda, where frequencies and proportion tables are highlighted to describe characteristics of socio-demographic with different teenage pregnancy.

Predict the relationship between socio-demographic characteristics and teenage pregnancy in Rwanda Non-parametric chi-square test (test of independence) will be used to measure the association between independent variables and dependent variables.

A logistic regression model that considers two variables will be employed to evaluate how each independent variable relates to teenage pregnancy. Thereafter, significant variables in the bivariate analysis will be considered for multivariable logistic regression to adjust for possible confounding and correlation among independent variables. Through the analysis, the p-value for each variable in logistic regression will be determined using the Wald test, and the level of statistical significance will be determined using a 95% confidence interval and a p-value of 0.05.

After conducting the inferential analysis, we are also planning in Predicting the association between socio-demographic factors and teenage pregnancy in Rwanda using a machine learning model that incorporates logistic regression, a decision tree classifier, and K Nearest Neighbor. and using these machine learning models to assess which best model performs better on our dataset. (Predict the number of teenage pregnancies.

3.5.2 Correlation and Chi square

In Data Science and Machine Learning, we frequently come across conditions where it is important for us to execute both feature selection and variable analysis. The correlation matrix helps to analyze the relationship between the data variables, it is represented between the range of 0 and 1 correlation value. Ideal positive correlation is represented by a value of 1, ideal negative correlation by a value of -1, and the absence of any relationship by a value of 0.

A correlation matrix is a table that displays the correlation coefficients that are present between a number of different variables. It is used to explore the relationships and dependencies that exist between the many variables contained in a dataset. The correlation matrix is used for a number of things, including selecting factors employed in dimensionality reduction methods such as principal component analysis (PCA) as well as identifying the correlations between variables in exploratory data analysis.

3.5.3 Categorical Variables

The chi-square test is a statistical method used to determine if there is a significant relationship via both categorical variables. It contrasts the observed frequencies in a contingency table with the frequencies that would be anticipated if the variables were assumed to be independent. When the chi-square test yields a significant result, the variables are likely to be related.

3.5.4 Confusion Matrix

The Matrix of Confusion The effectiveness of a classification method can be summarized in a table called a confusion matrix. A classification model's accuracy in producing predictions can be measured by examining its confusion matrix values. False positive (FP), false negative (FN), and true positive (TP) rates are listed in a table called the confusion matrix that predict to summarize the performance of a classification model as show in the figure 1.

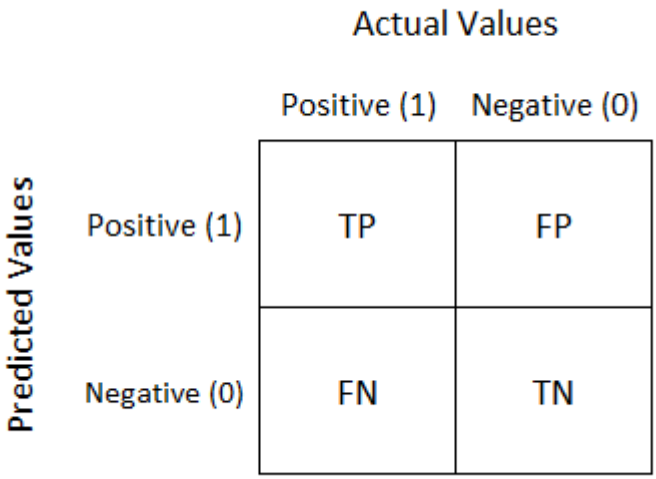

Figure 1 Confusion Matrix

i. Accuracy is calculated as a measure of how well the model makes predictions in general, we use equation 1.

$$Acc = \frac{TP+TN}{TP+TN+FP+FN}$$ **Equation 1**

ii. Recall is a metric used to assess how well a model does at identifying positive data. It is related to sensitivity and true positive rate. To determine this, we use equation 2

$$\frac{TP}{TP+FN}$$ **Equation 2**

iii. Specificity, or real negative rate, is a metric for assessing how well a model does at identifying true negative events, to determine this we use equation 3

$$\frac{TN}{TN+FP}$$ **Equation 3**

IV. RESULTS AND DISCUSSION

Some of the factors were found to be linked with teenage pregnancies like the teen behavior factor (Age at first sex, number of sex partners, the frequency of sexual intercourse), family background of the teen where many of teenage who become pregnant are from low educated parents, parents in the lowest quintile, parents whose marital status is not legally recognized, The size of the family (Household), etc.

Another factor that has been identified was linked to the knowledge and information that teen had on their reproductive health whereby teen do not know the use of family planning or did not hear about Family planning either on Radio, TV or in newspaper.

4.1 Explanatory statistic

Generally, teenage pregnancies were estimated at 7.3 percent in 2014/2015 and have declined to 5.3 percent in 2019/2020 with a combination rate of 6.2 percent. Several factors are associated with teenage pregnancies, and the study has grouped those factors into socio-demographic factors, Family factors, Health information factors, and reproductive factors as below:

Socio-demographic factors:

- There are a greater number of variables that are correlated to the dependent variable than there are variables that are not connected. The correction between socio demographic variables with teenage pregnancy show that the Region (Province) is highly correlated with teenage pregnancy as well as age of the teen with enough statistical significance. Teens residing in the southern provinces of Rwanda have a higher probability of experiencing teenage pregnancy in comparison to their counterparts residing in other provinces of the country, but also the study did not have enough evidence to conclude the statistically significant
- The Study found that teens in age 16- 19 are most affected by teenage pregnancies with a rate of 14 % with a significant relationship with teenage pregnancies with a p value equal to 0.002. The prevalence for teenage pregnancies was found at 6.2% and was highly influenced with sexual activity of the teens where 0% of teens have had their 1st sex at the age under 10years and 10.6% of teens in age range from 10 -14 years and a 42,8% in the age range from 15-19 years
- Education of teens was also highly correlated with an inverse relationship with teenage pregnancies; as far as the teen is educated the lower is the rate of teenage pregnancy. Furthermore, education level of the teen also has shown strong association with teenage pregnancy. Surprisingly, teens that have primary level of education was found to have high rate of pregnancy as compared with those with higher level of education. Teens who are not educated are 17.2 percent higher to the ones who have primary level 8.3 percent and higher on those with secondary level of education 3.3 percent.

Behavior factors:

- we found that having multiple sex partners were associated with teenage pregnancies with P-value Less than 0.006. It has been found that teen with more than one partner are most affected with pregnancy as compares to those with one sex partner. For example, the prevalence of teenage pregnancies was 80.2 among those who have 5 or more sex partners as compared to 31.6 for those with 1sex partner. Other variables that were associated with teenage pregnancies we can mention number of frequencies of having sex and the use of contraceptives.

Family background factor:

- Characteristics like the wealth index of the family, the household size was also linked with teen pregnancies with P-value below 0,05. this point that family wealth and household size are factors that may contribute to the occurrence of teenage pregnancies. It suggests that teenage from lower socioeconomic backgrounds or those living in larger households may be more susceptible to experiencing teenage pregnancies

Access to health Information factors:

- Teenagers who lack access to radio health information have a higher likelihood of being affected by teenage pregnancies, as opposed to their counterparts who have access to various forms of media. The availability of information plays a significant role in influencing teenage pregnancy rates, and those adolescents who do not engage with radio health information are particularly vulnerable to this issue.

Sexual reproductive factors:

- Some factors were classified as sexual reproductive including the knowledge on family planning, The proximity to the healthcare facility and instances of domestic violence among teenagers are inversely related. The study surprisingly found out that teen who declared to have knowledge about family planning were most affected with pregnancy with 6.6 percent as compared to 0.5 among those with no knowledge about family planning method. Teen who was sexual violated were most affected with teen pregnancy with 13.9 percent and those who had physical violence with 11.8 percent.

4.2 Chi-Square test

Chi-square test results between the independent variable and teenage pregnancy.

Figure 2 Chi Square

If the null hypothesis is correct, p-value reflects the likelihood of observing the test statistic (or a more extreme number). In the chi-square test, the null hypothesis says that there is no correlation between the two factors. The observed link is likely statistically significant if the p-value is less than a preset significance level, such as 0.05, in which case the null hypothesis can be rejected in favor of the alternative hypothesis. Based on the information presented in Figure 2 above, it can be observed that a larger number of independent variables (socio-demographic factors, Family factors, Health information factors, and reproductive factors) exhibit correlations with the dependent variable (teenage pregnancy) compared to the number of variables that show no connection.
After understanding figure 2, Correction between socio-demographic variables and teen pregnancy shows that our study found significant associations between teen pregnancy and socio-demographic factors in Rwanda, and has found statistical relationship among different variable, therefore contradicted the hypothesis (H0) that teen pregnancy is not associated to socio-demographic characteristics and verified our hypothesis. We may conclude that there is a statistical correlation between teen pregnancies and socio-demographic characteristics as well as other factors.

In terms of education level, we identified a strong inverse correlation between educational attainment and teenage pregnancy. We discovered a significant negative correlation between the level of education and the occurrence of teenage pregnancy. Teenagers who had a lesser extent of education had a higher probability of getting pregnant in contrast to those who achieved higher levels of education. This finding highlights the importance of education in decreasing the risk of teenager's pregnancies.

In terms of region, we explored the association between region and teenage pregnancy rates. After conducting our statistical analysis, the calculated p-value of 0.14 ($p > 0.05$) indicates that the level of statistical significance is not sufficiently strong. This p-value suggests that there may not be a substantial relationship between the variables under investigation. This means that the association between region and teenage pregnancy is not strong enough to conclude with high confidence that the association is statistically significant.

Age categories has a P-values less than 0.05 therefore we reject the null hypothesis and accept the alternative hypothesis which mean that there is correction between age categories and teenage pregnancy null hypothesis therefore we fail to reject the alternative hypothesis.

4.3 Comparison of Models

Table 1 below shows the K-Nearest Neighbor (KNN) appears to be the best among the three models. Here's why:

- Accuracy: The K-nearest neighbors (KNN) model attains the utmost level of precision of 95.86%, indicating that it correctly predicts the outcome in 95.86% of cases.
- Specificity: The KNN model also has the highest specificity of 58.33%. Specificity represents the proportion of correctly identified negative cases (non-teenage pregnancy) out of all the actual negative cases. A higher specificity indicates a better ability to correctly identify non-teenage pregnancy cases.
- Recall: The KNN model achieves a recall of 98%, which is the highest among the three models. Recall represents the proportion of correctly identified positive cases (teenage pregnancy) out of all the actual positive cases. A higher recall indicates a better ability to correctly identify teenage pregnancy cases.

Considering these metrics, the KNN model outperforms Logistic Regression and Decision Tree in terms of accuracy, specificity, and recall. It demonstrates a better overall performance in predicting teenage pregnancy.

Table 1: Comparison of the models

Models	Accuracy	Specificity	Recall
K-Nearest Neighbor	95,86%	58,33%	98%
Logistic Regression	95,07%	51,16%	97%
Decision Tree	95,07%	51,16%	97%

V. ACKNOWLEDGMENTS

I am extremely thankful to everyone who has played a part in the successful culmination of this thesis entitled the Prediction of teenage pregnancy in Rwanda.

Above all, I express my gratitude to my mentor, Dr. NIYIZAMWIYITIRA Christine, for their immeasurable guidance, support, and expertise throughout the entire research process. Their insightful feedback, constructive feedback, and unwavering encouragement have been instrumental in shaping this study.

I am immensely grateful to the participants who willingly shared their time and personal experiences, allowing me to share with me the necessary data for this research. Their openness and cooperation have significantly contributed to the findings' validity and relevance.

I would like to acknowledge my friends and family's support and encouragement throughout this academic journey. Their unwavering belief in my abilities and their constant motivation has strengthened me during challenging times.

I would also like to acknowledge the contributions of the researchers and scholars whose works have been referenced in this dissertation. Their groundbreaking studies and insightful publications have provided the necessary foundation and inspiration for this research.

Lastly, I would like to extend my appreciation to both the academic community and the African Center of Excellence in Data Science for providing an enriching and stimulating environment for learning and research. The resources, facilities, and opportunities provided have played a significant role in the successful completion of this dissertation.

In conclusion, I am humbled and grateful to all my family, and friends, in the realization of this research and this journey of 3 years in ACE-DS. I appreciate your unwavering support and all the kindness you have shown me.

REFERENCES

Adogu, P., Udigwe, I., Udigwe, G., & Ubajaka, C. (2014). Review of problems of adolescent sexual behavior and the role of Millennium Development Goals 4, 5, and 6 in Nigeria. *International Journal of Clinical Medicine, 5(15).*

Akella, D., & Jordan, M. (2015). Understanding teenage pregnancy in a contemporary society. *Journal of Adolescent Health, 57(3), 249-254.*

Brace, A. M., Hall, M., Eatman, J., & Tucker, N. (2008). Social, Economic and Health Costs of Unintended Teen Pregnancy: The Circle of Care Intervention Program in Troup County, Georgia. *Journal of the Georgia Public Health Association, 1(1), 33-46.*

Brown, B. B., Mounts, N., Lamborn, S. D., & Steinberg, L. (1993). Parenting practices and peer group affiliation in adolescence. Child Development, 64(2), 467–482.

Domenico, D. M., & Jones, K. H. (2007). Adolescent pregnancy in America: Causes and responses. *Journal for Vocational Special Needs Education*, 30(1), 4–12.

Hakizimana, E. (2019). Kirehe District counts 600 teenage pregnancies, victims still hide the culprits. Rwanda Inspirer.

Hamilton, B. E., Martin, J. A., & Ventura, S. J. (2009). Births: Preliminary data for 2007. National Vital Statistics Reports, 57(12), 1-23.

Huseyin Naci, & Ioannidis, J. P. A. (2015). Evaluation of wellness determinants and interventions by citizen scientists. JAMA, 314(2), 121-122.

Ikramullah, E., Manlove, J., Cui, C., & Moore, K. A. (2009). Parents matter: The role of parents in teens' decisions about sex. Child Trends Research Brief, 45, 1-7.

Irin. (2007). Lack of family support hampers youth reproductive health. Retrieved from https://www.irinnews.org/report/72697/zimbabwe-lack-family-support-hampers-youth-reproductive-health

Islam, M. M. (2022). Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda. BMC Pregnancy Childbirth, 22- 348.

Iwu, E. N., Onoja, A. J., Ijioma, J. C., Ngumah, E. C., & Egeruoh, M. C. (2011). Sociocultural determinants of teenage pregnancy in Nigeria. Cross-Cultural Communication, 7(2), 129-137.

Iyakaremye, I. (2018). The readiness of the community to address teenage pregnancy in Rwanda was a study conducted in the seven districts of Eastern Province. Haguruka Report.

Jaccard, J., Blanton, H., & Dodge, T. (2005). Peer influences on risk behavior: An analysis of the effects of a close friend. Developmental Psychology, 41(1), 135-147.

Kimemia, M. (2015). Influence of cultural beliefs, social and financial factors, internet use, and peer pressure on teenagers' pregnancies in schools: A Kenyan perspective. *International Journal of Education and Research*, 3(7), 145-156.

World Health Organization. (2014). Adolescent pregnancy: Unmet needs and undone deeds - A review of the literature and programmes. BJOG: An International *Journal of Obstetrics and Gynaecology*, 121(Suppl 1), 64-72.

Ybarra, M. L., Espelage, D. L., & Mitchell, K. J. (2007). The co-occurrence of Internet harassement and unwanted sexual sollicitation victimisation and perpetration: Associations with psychosocial indicators. *Journal Adolescent Health*, 7(9), 342.